Creating an early diagnostic test for Alzheimer's Disease

Voyle, Nicola Joanne

Awarding institution:
King's College London

Creating an early diagnostic test for Alzheimer's Disease

Nicola Joanne Voyle

Social, Genetic and Developmental Psychiatry Centre

Institute of Psychiatry, Psychology and Neuroscience

King's College London

A thesis submitted to King's College London

for the degree of Doctor of Philosophy

July 2016

Acknowledgements

Firstly, I would like to thank my supervisors Dr. Steven Kiddle and Dr. Richard Dobson for their priceless support, advice and tolerance of my nagging. I am also grateful for the help I have received from many of the Bioinformatics group at the SGDP Centre.

I wouldn't have been able to finish this work without the love, support and encouragement of my family. Thank you mum and dad for believing in me, mopping up my tears and for all your words of wisdom. Thank you to my sister Sam for keeping my feet on the ground and for never failing to provide a good giggle. And finally, thank you Joe. Thank you for your patience, encouragement and belief in me - I'll finally stop being a student now!

This thesis would never have been finished without you all.

Abstract

The aim of the research in this thesis was to discover and validate blood biomarkers of early Alzheimer's Disease (AD). Existing and novel datasets from cohort studies were used for discovery and to attempt validation of previously reported biomarkers. For example, this thesis presents the first study to investigate associations between brain amyloid and blood metabolites. Further, this thesis presents the first study to combine more than one modality of blood biomarker in AD research and the first study to use a Bayesian methodology in this field.

This thesis begins by aiming to validate candidate protein markers of brain amyloid burden in a novel proteomics dataset. Secondly, pathway-based methods are used to investigate the use of gene expression measurements as a potential biomarker of AD diagnosis. In the fourth chapter I generated a novel metabolomics dataset to investigate associations between blood metabolites and brain amyloid burden. A panel is found that predicts dichotomized amyloid burden with reasonable accuracy. The accuracy is improved by the inclusion of a candidate protein in the model.

The fifth chapter of this thesis is focused on the use of a Bayesian methodology to predict measurements of amyloid using a variety of *omics* data. The Bayesian methodology allows incorporation of historical information by placing informative priors on demographic variables. No improvement is seen over demographics alone. The final chapter of this thesis aims to predict amyloid and tau burden using a polygenic risk score and levels of tau in blood. I have also considered a combined amyloid and tau pathology endpoint. The blood markers considered here do not improve predictive ability over demographics alone.

Much of the work in this thesis highlights the importance of demographic factors in the diagnosis of early AD. The metabolite discovery work shows an improvement

in predictive ability over demographics alone and warrants further investigation and replication. The other chapters of this thesis highlight that (in the settings investigated so far) blood measurements add minimal information above demographics alone.

Statement of Authorship

All work included in this thesis, both analysis and written, was performed by Nicola Voyle with the following exceptions:

1. All published papers were circulated for co-author review and subject to peer review prior to publication. Manuscripts, and occasionally analyses, were updated subject to these reviews.

2. The datasets used in all publications were collected by others. In all cases this is made clear in the methodology sections of each chapter.

This work was generously funded by the Alzheimer's Society.

List of Abbreviations

Abbreviation	Details
2DGE	2-dimensional gel electrophoresis
Aβ	Amyloid-β
AchE	Acetylcholinesterase
AchEIs	Acetylcholinesterase Inhibitors
AD	Alzheimer's Disease
ADAS-COG	Alzheimer's Disease Assessment Scale - Cognitive Subscale
ADNI	Alzheimer's Disease Neuroimaging Initiative
ADSP	Alzheimer's Disease Sequencing Project
AIBL	Australian Imaging, Biomarker and Lifestyle Flagship Study of Ageing
AIC	Akaike's Information Criteria
ANM	AddNeuroMed
API	Alzheimer's Prevention Initiative
APOE	Apolipoprotein E
APP	Amyloid precursor protein
ARUK	Alzheimer's Research UK
AUC	Area Under the Curve
BBB	Blood-brain barrier
BRC	Biomedical Research Centre
CANTAB-PAL	Cambridge Neuropsychological Test Automated Battery - Paired Associate Learning
CDR	Clinical Dementia Rating
CI	Confidence Interval

CRF	Conditional Random Forest
CSF	Cerebrospinal Fluid
CTL	Control
CV	Cross validation
DAD	Disability Assessment for Dementia
DCR	Dementia Case Registry
DESCRIPA	Development of screening guidelines and criteria for pre-dementia Alzheimer's disease
DNA	Deoxyribonucleic Acid
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
EDAR	Beta amyloid oligomers in the early diagnosis of AD and as marker for treatment response
EMIF	European Medical Information Framework
EOAD	Early Onset Alzheimer's Disease
EPAD	European Prevention of Alzheimer's Dementia consortium
FAD	Familial Alzheimer's Disease
FDA	United States Food and Drug Administration
FGG	Fibrinogen Gamma Chain
fMRI	Functional Magnetic Resonance Imaging
FTD	Frontotemporal Dementia
GEE	Generalized Estimating Equation
GERS	Gene Expression Risk Score
GSVA	Gene Set Variation Analysis
GUI	Graphical User Interface
GWAS	Genome Wide Association Studies
HC	Healthy Control
IGAP	International Genomics of Alzheimer's Project
IgM	Immunoglobulin M
IQR	Inter-quartile Range
IWG	International Working Group

LC-MS/MS	Liquid Chromatography Tandem Mass Spectrometry
LD	Linkage Disequilibrium
LMG	Lindemann, Merenda and Gold
MAPT	Microtubule-associated protein tau
MCI	Mild Cognitive Impairment
MCMC	Markov Chain Monte Carlo
MMSE	Mini Mental State Exam
MRI	Magnetic Resonance Imaging
mRNA	Messenger Ribonucleic Acid
NAB	Neocortical Amyloid Burden
NHGRI	National Human Genome Research Institute
NICE	National Institute of Health and Care Excellence
NMDA	N-methyl-D-aspartate
OOB	Out of bag
OR	Odds ratio
PC1	First Principal Component
PCA	Principal Components Analysis
PET	Positron Emission Tomography
PGRS	Polygenic Risk Score
PHF	Paired Helical Filaments
PC	Phosphatidylcholine
PE	Phosphatidylethanolamine
PiB	Pittsburgh Compound B
PLAGE	Pathway Level Analysis of Gene Expression
PLS	Partial Least Squares
PPY	Pancreatic Polypeptide
PSEN1	Presenilin 1
PSEN2	Presenilin 2
QC	Quality Control
RBM MAP	Rules Based Medicine Multi-analytes Profile

RF	Random Forests
RFE	Recursive Feature Elimination
RMSE	Root Mean Squared Error
RNA	Ribonucleic Acid
ROC	Receiver Operator Characteristic
SD	Standard Deviation
SE	Standard Error
SGDP	Social, Genetic and Developmental Psychiatry
SIMOA	Single Molecule Array
SNP	Single Nucleotide Polymorphism
SOMA	Slow Off-Rate Modified Aptamer
SPM	Statistical Parametric Mapping
SSGSEA	Single Sample Gene Set Enrichment Analysis
SUVR	Standardized Uptake Value Ratio
SVM	Support Vector Machine
TREM2	Triggering Receptor Expressed on Myeloid Cells 2
UCSF	University of California, San Francisco
UK	United Kingdom
US	United States
USA	United States of America
WGS	Whole Genome Sequencing

Contents

List of Figures

List of Publications

Cristina Menni, Steven J. Kiddle, Massimo Mangino, Ana Viñuela, Maria Psatha, Claire Steves, Martina Sattlecker, Alfonso Buil, Stephen Newhouse, Sally Nelson, Stephen Williams, **Nicola Voyle**, Hilkka Soininen, Iwona Kloszewska, Patrizia Mecocci, Magda Tsolaki, Bruno Vellas, Simon Lovestone, Tim D. Spector, Richard J. Dobson, and Ana M. Valdes (2015) Circulating Proteomic Signatures of Chronological Age. *Journals of Gerontology, Series A: Biological and medical sciences*, 809 - 816.

Nicola Voyle, David Baker, Samantha C. Burnham, Antonia Covin, Zhanpan Zhang, Dipen P. Sangurdekar, Cristina A. Tan Hehir, Chantal Bazenet, Simon Lovestone, Steven Kiddle, Richard JB. Dobson and the AIBL research group (2015) Blood protein markers of Neocortical Amyloid-β Burden: A candidate study using SOMAscan technology. *Journal of Alzheimers Disease* 46 (2), 947 - 962.

Justin TW. Chiam, Katie Lunnon, **Nicola Voyle**, Petroula Proitsi, Giovanni Coppola, Daniel Geschwind, Sally Nelson, Caroline Johnston, Hilkka Soininen, Iwona Koszewska, Patrizia Mecocci, Magda Tsolaki, Bruno Vellas, Angela Hodges, Simon Lovestone, Stephen Newhouse, Richard JB. Dobson, Steven J. Kiddle and Martina Sattlecker (2015) No evidence to suggest that the use of acetylcholinesterase inhibitors confounds the results of two blood-based biomarker studies in Alzheimers disease. *Journal of Alzheimers Disease* 47 (3), 741 - 750.

Nicola Voyle, Aoife Keohane, Stephen Newhouse, Katie Lunnon, Caroline

Johnston, Hilkka Soininen, Iwona Kloszewska, Patrizia Mecocci, Magda Tsolaki, Bruno Vellas, Simon Lovestone, Angela Hodges, Steven J. Kiddle and Richard JB. Dobson on behalf of the AddNeuroMed consortium (2015) A pathway based classification method for analysing gene expression for Alzheimers Disease diagnosis. *Journal of Alzheimers Disease* 49 (3), 659 - 669.

Nicola Voyle, Min Kim, Petroula Proitsi, Nicholas J. Ashton, Alison L. Baird, Chantal Bazenet, Abdul Hye, Sarah Westwood, Raymond Chung, Malcolm Ward, Gil D. Rabinovici, Simon Lovestone, Gerome Breen, Cristina Legido-Quigley, Richard JB. Dobson and Steven J. Kiddle for the Alzheimer's Disease Neuroimaging Initiative (2016) Blood metabolite markers of Neocortical Amyloid Burden: Discovery and enrichment using candidate proteins. *Translational Psychiatry* 6, e719.

Nicola Voyle, Willemijn Jansen, Aoife Keohane, Hamel Patel, Amos Folarin, Stephen Newhouse, Caroline Johnston, Kuang Lin, Pieter Jelle Visser, Angela Hodges, Richard JB. Dobson and Steven J. Kiddle for the Alzheimer's Disease Neuroimaging Initiative (2016) Do peripheral markers help to predict amyloid burden in a non-demented population? A Bayesian approach. *Submitted to PLoS ONE.*

Nicola Voyle, Hamel Patel, Amos Folarin, Stephen Newhouse, Caroline Johnston, Pieter Jelle Visser, Richard JB. Dobson and Steven J. Kiddle for the Alzheimer's Disease Neuroimaging Initiative (2016) Genetic risk as a marker of amyloid-β and tau burden in cerebrospinal fluid. *Accepted at the Journal of Alzheimer's Disease.*

Chapter 1

Introduction

1.1 Alzheimer's Disease

Alzheimer's Disease (AD) was first described in 1906 by Alois Alzheimer (Maurer *et al.*, 1997). He was faced with the case of a 51 year old woman, Auguste D, whose symptoms included memory loss, disorientation, hallucinations and psycho-social incompetence. Alzheimer was unable to reach a diagnosis but was interested in her condition and monitored Auguste closely. When she passed away Alzheimer performed an autopsy that resulted in the first records of amyloid plaques in the brain. One hundred and ten years on, many advances have been made in understanding dementia, raising awareness and providing suitable care. However, advances in treatments for the disease have been few and far between (Corbett & Ballard, 2012).

In modern medicine, AD falls under the umbrella term of neurocognitive disorder formerly classified as dementia. Such disorders are characterised by a decline in cognitive functioning from a level previously attained. In particular, the Diagnostic and Statistical Manual of Mental Disorders Fifth Edition (DSM-5) defines six cognitive domains: perceptual-motor function, language, learning and memory, social cognition, complex attention and executive function. DSM-5 diagnosis of probable AD requires decline in at least two cognitive domains one of which should be memory and learning (Sachdev *et al.*, 2014). A variety of symptoms are associated with AD and there is large variation between individuals. Symptoms can include memory loss, confusion, hallucinations and sleep disturbances eventually leading to complete

loss of speech, movement and memory. Ferri *et al.* (2005) estimated the global dementia prevalence at 24.3 million people with 4.6 million new cases developing each year. With an ageing population the number of people who develop this condition, and the cost of their care, is rapidly increasing. AD, as the most common form of dementia, will pose a significant socio-economic burden to the society of the future (Wimo *et al.*, 2013).

There are two main types of AD: early onset and late onset (or sporadic) AD. Their combined prevalence in the United States (US) was estimated at 5.3 million people in 2015 with a projected incidence of 615,000 new cases a year by 2030 (Alzheimer's Association, 2015).

1.1.1 Early Onset Alzheimer's Disease

Early onset Alzheimer's Disease (EOAD) is a Mendelian disorder and is typically associated with an age of disease onset younger than 65 years. A Mendelian disorder is one where the presence of a mutation at a specific gene locus guarantees development of the disease. In the case of EOAD the cause of disease is an autosomal, dominant mutation in either the Presenilin 1 (*PSEN1*), Presenilin 2 (*PSEN2*) or amyloid precursor protein (*APP*) genes. Although the development of disease is a given if at least one of these gene mutations is present, several diagnoses could be made under the International Working Group (IWG-2) criteria (Dubois *et al.*, 2014). The gene mutations provide evidence of *in-vivo* AD pathology for pre-clinical states of AD, typical AD and atypical AD.

In the US it is believed that around 200,000 people have a diagnosis of EOAD (Alzheimer's Association, 2015). This is approximately one 25th of the number of people who have the late onset form of AD.

1.1.2 Late onset Alzheimer's Disease

LOAD has a much higher prevalence with an estimated 5.1 million people living with the disease in the US (Alzheimer's Association, 2015). LOAD has a complex etiology and is consequently less well understood than EOAD. Presence of LOAD

cannot be completely determined by one genetic locus or group of loci. It is thought that the disease process is controlled by a range of genetic and environmental risk factors each conveying varying levels of risk for disease.

The remainder of this thesis is focused on LOAD, which for brevity will be referred to as AD from here onwards.

1.1.3 Disease pathology

It has been hypothesised that the cause of AD is a build up of amyloid-β (Aβ) plaques and hyperphosphorylated tau tangles in the brain. The senile plaques in the brains of people with AD were first observed by Alzheimer himself in 1906. It wasn't until 1984 that Glenner & Wong (1984) identified the plaques to be made of Aβ protein. Amyloids are insoluble, mis-folded proteins that have been associated with several diseases including Huntington's Disease, Rheumatoid Arthritis and Amyloidosis (Ramirez-Alvarado *et al.*, 2000). In particular, Aβ is most commonly found in its 40 and 42 amino acid forms as a result of cleavage of the amyloid precursor protein (APP) by α and γ secretase. The normal function of Aβ is not well understood but Aβ abundances in the brain have been observed to increase during ageing (Hiltunen *et al.*, 2009). Tau proteins are largely found in neuronal cells with their primary function being to stabilize microtubules. They are coded for by the microtubule-associated protein tau ($MAPT$) gene, found on chromosome 17, and when hyperphosphorylated form pathogenic tangles. Tau proteins have been associated with other neurodegenerative disorders besides AD, including frontotemporal dementia (FTD), and are also present in high levels in the cerebrospinal fluid (CSF) of people who have suffered a traumatic brain injury (Ling *et al.*, 2015).

It is extremely difficult to directly study the causes of AD. Several factors contribute to this, most obviously the inaccessibility of brain tissue in a living human. Post-mortem studies can be conducted on brain samples but the tissue is incredibly complex and slow to develop. Furthermore, the heterogeneity of pathology between individuals is very high. This means large sample sizes are often needed to perform any meaningful analysis. Bodily fluids such as CSF, blood, urine and saliva are far

more accessible. However, other than CSF, they are separated from the brain by the blood-brain barrier (BBB). As this is only a semi-permeable membrane, the chances of signal passing from the brain into surrounding fluids are reduced. Although, it has been observed that the BBB weakens with age which may aid us in detecting a signal in blood (Montagne *et al.*, 2015).

The amyloid cascade hypothesis (Figure 1.1) is a common theory detailing how the aggregation of Aβ and tau results in AD. It states that Aβ plaques precede tau tangles and start the process of tau accumulation. The hypothesis states that, given enough time, the development of this pathology will always lead to AD through damage to synapses and eventually neurons (Karran *et al.*, 2011). The popular theoretical model of disease progression suggested by Jack *et al.* (2013) details our current understanding of the sequential changes in Aβ, tau, brain structure, memory and cognitive function over time. This model is consistent with the amyloid cascade hypothesis as it suggests that the aggregation of Aβ plaques precedes the presence of other AD pathologies. Other evidence in support of the amyloid cascade hypothesis includes the fact that Down's syndrome greatly increases an individuals risk of developing AD (Wiseman, 2015). Down's syndrome is caused by trisomy of chromosome 21, the same chromosome that holds the *APP* gene.

There are some fundamental questions that challenge the amyloid cascade hypothesis (Morris *et al.*, 2014). In particular, many people with high Aβ burden at autopsy show no signs of cognitive impairment during their lifetime. It has been suggested that this is because they did not live long enough for the symptoms of AD to develop. This may be true but is impossible to verify. Further, the amyloid cascade hypothesis has so far failed to provide successes in drug development and the genetic hits associated with AD suggest a more complex disease mechanism (Lambert *et al.*, 2013). Nevertheless, it is clear that somehow Aβ is involved in the etiology of AD. This is even highlighted in theories opposing the amyloid cascade hypothesis with the main difference being the inclusion of inflammation or tau as the primary initiator of Aβ accumulation (Goedert, 2004; Morris *et al.*, 2013).

The theoretical model of Jack *et al.* (2013) also covers the concept of mild cog-

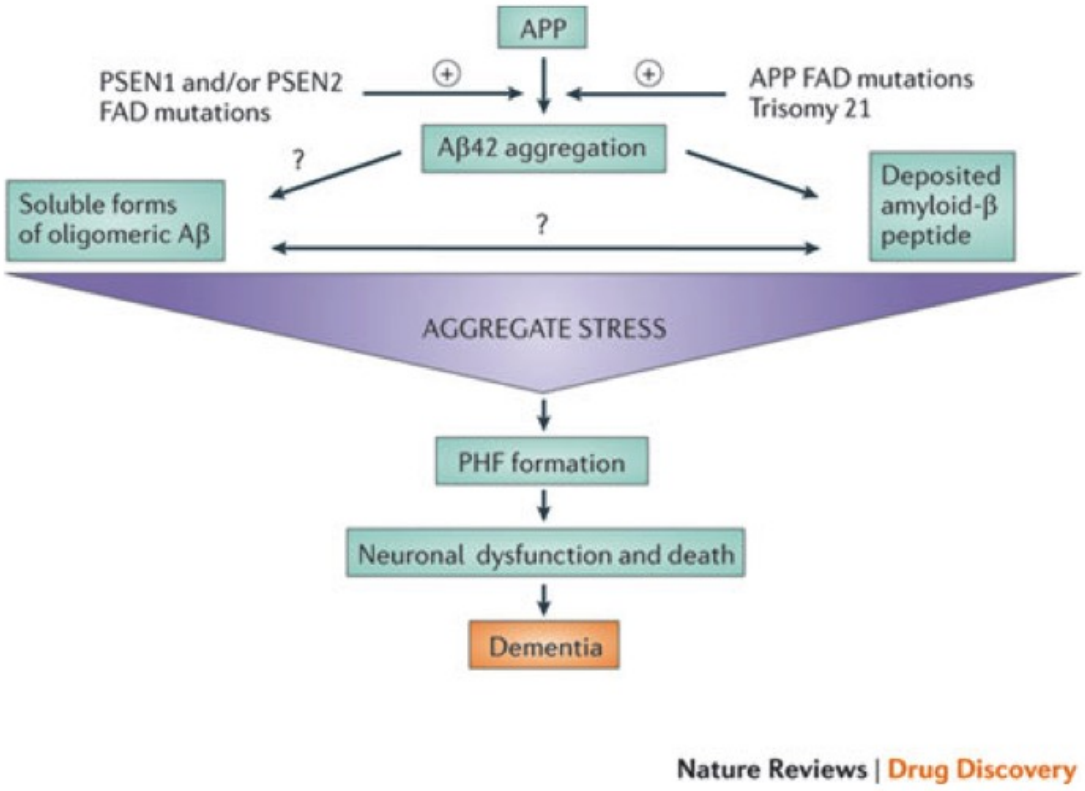


Figure 1.1: The amyloid cascade hypothesis, reproduced from Karran *et al.* (2011). *FAD = Familial Alzheimer's Disease; PHF = Paired Helical Filaments.*

nitive impairment (MCI), as a precursor to AD. MCI is characterised as a minor cognitive deficit particularly seen in the intermediate stages between the normal cognitive decline of ageing and the quicker, more severe, decline associated with dementia. Although it can be a precursor for developing AD, MCI can also remain stable or lead to other dementias. It has been proposed that this is because MCI can be caused by a range of underlying pathologies; MCI individuals with increased Aβ and tau burdens are most likely to progress to AD dementia (Okello *et al.*, 2009).

1.1.4 Risk factors for Alzheimer's Disease

AD has been identified as a complex disease meaning its presence or absence is not completely determined by one genetic locus or set of loci. It is instead caused by a variety of environmental and genetic factors.

1.1.4.1 Non-genetic risk factors of Alzheimer's Disease

Age is the largest risk factor for LOAD; as an individual gets older their risk of developing the disease increases substantially up to a doubling in age-specific prevalence every 5 years after age 65 (See Figure 1.2) (Alzheimer's Association, 2015; Qiu *et al.*, 2009). Further, epidemiological factors also drive AD risk with people who have lower levels of education having a higher risk of AD (Räihä *et al.*, 1998). It has been hypothesised that this may be due to 'brain reserve' generated through elongated periods of study. As with many diseases, AD has associations with common conditions such as obesity, cardiovascular disease and diabetes. Cognitive decline in AD is thought to be faster in individuals living alone. Norton *et al.* (2014) have studied the population attributable risk for AD of 7 modifiable risk factors: midlife obesity and hypertension, physical inactivity, diabetes, smoking, low educational attainment and depression. They concluded that 28.2% of AD cases worldwide can be attributed to these risk factors. Matthews *et al.* (2016) have recently shown a reduced incidence of AD in British males over 65 years thought to be driven by changes in these modifiable risk factors.

1.1.4.2 Genetic risk factors of Alzheimer's Disease

Heritability estimates for AD from twin studies are approximately 50-70% (Gatz *et al.*, 1997; Pedersen *et al.*, 2004). These studies provide upper-bounds for the proportion of variability one could expect to explain using genetics. AD is more common in females than males; some of this risk can be attributed to the longer life expectancy of females (Viña & Lloret, 2010).

The most well-researched genetic risk factor for AD is the apolipoprotein E (*APOE*) gene on chromosome 19. In the central nervous system the main function of the protein coded for by the *APOE* gene is to transport cholesterol and fat soluble vitamins and minerals to the neurons. There are three main allelic variants of *APOE* defined by two single nucleotide polymorphisms (SNPs): rs429358 and rs7412. The most common *APOE* variant is $\epsilon 3$ making the most common genotype $\epsilon 3 \epsilon 3$. The risk allele for AD is $\epsilon 4$. In comparison to the most common *APOE* genotype, one

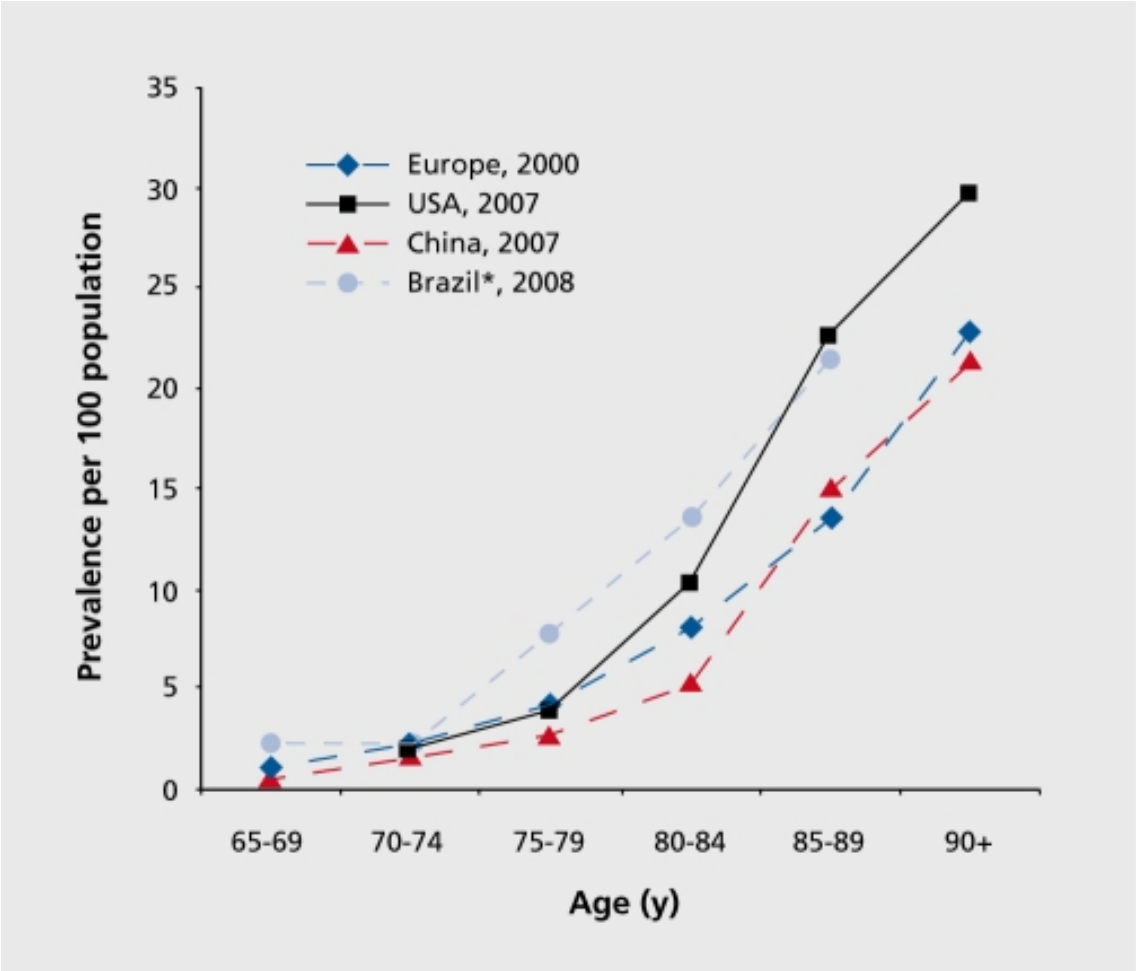


Figure 1.2: Age-specific AD prevalence (per 100 population) across continents and countries. * = prevalence of all types of dementia. Reproduced from Qiu *et al.* (2009).

copy of the risk allele in Caucasians leads to an odds ratio (OR) of 2.6 ($\epsilon 2\epsilon 4$) or 3.2 ($\epsilon 3\epsilon 4$). Homozygosity of the $\epsilon 4$ allele results in an OR of 14.9 (Liu *et al.*, 2013). Overall, *APOE* is estimated to explain approximately 4-6% of phenotypic variability (Ridge *et al.*, 2013; Lee *et al.*, 2013). Genetic risk factors are further discussed in section 1.2.2.

The heritability of the protein pathologies of AD: Aβ and tau, is less well researched. One small twin study (N = 17 twin pairs) found that cognitively normal co-twins with cognitively impaired probands showed an increased Aβ burden. The estimated increase of 117-121% indicates that the accumulation of Aβ pathology is influenced by genetic factors (Scheinin *et al.*, 2011).

1.1.5 Treatments for Alzheimer's Disease

There are a handful of drugs available to treat the symptoms of AD. The National Institute for Health and Care Excellence (NICE) guidelines (2011) for treatment of dementia recommend the use of Acetylcholinesterase (AchE) Inhibitors (AchEIs) for management of mild and moderate AD. These drugs block AchE from catalysing the break down of some neurotransmitters. For patients who cannot take AchEIs, or for whom the AD diagnosis is severe, memantine can be prescribed. Memantine blocks N-methyl-D-aspartate (NMDA) receptors stopping excess activity of the neurotransmitter glutamate. However, both these treatment options are only effective in a subset of people and even then their effect is often moderate, delaying disease progression by around 6-12 months (Rockwood, 2004).

There are currently no approved disease-modifying treatments for AD, despite 413 AD clinical trials being performed world-wide by 2012 (Cummings *et al.*, 2014). The majority of these trials were in symptomatic individuals. For example, Bapineuzumab is a humanized anti-Aβ antibody developed by Pfizer and Johnson and Johnson. Two phase III trials of Bapineuzumab did not meet their primary endpoints; significant difference from placebo in change from baseline in the Alzheimer's Disease Assessment Scale - Cognitive Subscale (ADAS-COG) and the Disability Assessment for Dementia (DAD) (Salloway *et al.*, 2014). However, it was noted that approximately 14% of participants in these studies had low Aβ burden. It was therefore hypothesised that enrichment for participants with high Aβ should be performed.

A second theory for the lack of successful treatments is that trials are targeting people who are already showing the characteristic symptoms of AD. It is known that AD pathology develops over a period of around 20 years meaning at the time symptoms appear the brain is already damaged, perhaps irreversibly so (Villemagne *et al.*, 2011; Jack *et al.*, 2013). A solution to this would be to recruit asymptomatic individuals into preventative trials. This idea motivates the focus of this thesis on early AD where possible. For this approach to be most effective only people who are at high risk of developing AD should be enrolled in studies, again suggesting

the use of Aβ enrichment in recruitment. In the preventative trial setting the use of high Aβ burden as an inclusion criterion is justified by its presence as the earliest known pathological change in the disease mechanism of AD (Jack *et al.*, 2013). Additionally, Aβ is the target of the primary class of novel treatments currently being developed and tested. Aβ enrichment is being used in the A4 prevention study and in novel, adaptive trials such as the European Prevention of Alzheimer's Dementia consortium (EPAD; www.ep-ad.org) (Sperling *et al.*, 2014). A further potential enrichment mechanism is the use of genetics. Enriching prevention trials using genetic biomarkers has previously been shown to be effective for Acute Macular Degeneration and Type I diabetes (Hu *et al.*, 2013). The *APOE* gene is a well known risk factor for AD and could consequently be used to enrich for the likelihood of developing AD by including participants carrying the $\epsilon 4$ risk allele. This approach is being pursued in the TOMMORROW (www.tommorrowstudy.com) and Alzheimer's Prevention Initiative (API) trials.

The primary focus of this thesis is to find biomarkers of Aβ. However, several disease mechanisms, alongside inhibition and clearance of Aβ, are also being investigated as potential drug targets. Bapineuzumab, mentioned above, follows an immunotherapy approach to clearing Aβ; an exciting area with over 10 agents entering clinical trials. Other approaches include targeting anti-inflammatory processes, accumulation of tau, *APOE* and investigation of metabolic dysfunction informed by discoveries in epidemiology and associations between AD and diabetes (Citron, 2010).

A slightly different approach to drug discovery in AD is the re-purposing of treatments already approved, or developed and abandoned, in different disease indications (Appleby *et al.*, 2013). Candidate agents are chosen based on relevant disease mechanisms. This method of re-purposing has several benefits over developing novel treatments. In particular, toxicology, safety and pharmacodynamic profiles, as well as manufacturing procedures, are well established leading to much shorter development times. An example of a promising re-purposed agent is metformin; a drug used to treat type II diabetes (Appleby *et al.*, 2013). Human trials have shown mixed

results with increased levels of Aβ observed when metformin is administered alone but decreased levels of Aβ when it is given in combination with insulin.

1.2 Biomarkers of AD and AD pathology

Symptoms of AD partially overlap with other dementias, meaning that a definitive diagnosis can only be made post-mortem. An accurate diagnosis is especially hard in early disease when the symptoms, such as mild memory loss and depression, are particularly non-specific. Both of these factors lead to misdiagnosis of patients. To protect against misdiagnosis it has been suggested that AD specific biomarker positivity should be included in diagnostic criteria (Dubois *et al.*, 2014). The US Food and Drugs Administration (FDA) define a biomarker as a 'laboratory measure that reflects the activity of a disease process' (Katz, 2004). This section introduces some key biomarkers of AD.

1.2.1 Non-Blood Biomarkers

1.2.1.1 Imaging

AD is a disease of the brain and consequently brain imaging is an obvious place to start in the search for a biomarker. Many studies have performed structural magnetic resonance imaging (MRI) scans on people with a diagnosis of probable AD. They have observed overall brain atrophy and, in particular, a shrinking of the hippocampus (the region of the brain associated with memory) and medial temporal lobe (Kehoe *et al.*, 2014). Structural MRI measures have been used in cohort studies to predict conversion from MCI to AD (Westman *et al.*, 2011). Meanwhile, newer techniques such as functional MRI are currently being validated as diagnostic tools for AD; fMRI has the ability to measure changes in the brain by monitoring blood flow. Although useful, MRI scans seem unable to pick up subtle changes in brain composition before symptoms arise so provide limited information for diagnosis of early AD. Furthermore, MRI scans are useful for visualising brain structures but are not pathology specific; they do not give an indication of the levels or location of

tau and Aβ (Johnson *et al.*, 2012).

To this end, ligands of the Aβ and tau proteins have been developed for use in conjunction with positron emission tomography (PET) imaging. The ligands give an impression of the location and quantity of pathology as measured by fluorescence and are proving very useful measures in endophenotypic studies of AD. In recent studies the approved ligands have shown good, but not perfect, correlation with plaques and tangles at autopsy particularly when using a dichotomized measure of pathology burden (Clark *et al.*, 2012; Ariza *et al.*, 2015).

1.2.1.2 CSF

CSF is a clear liquid that surrounds the brain and central nervous system. It protects the brain and spinal chord from damage while performing waste clearance and homoeostasis. This role makes it a prime candidate to contain substances informative of the presence of any brain pathology. Studies have shown that Aβ burden in the brain is strongly, negatively correlated with Aβ in CSF (Landau *et al.*, 2013). On the other hand, brain tau burden is positively correlated with tau in the CSF. Due to the imperfect correlation of Aβ and tau between brain and CSF, researchers have searched a plenitude of other candidate molecules for markers of AD. For example, Zürbig & Jahn (2012) review proteomics in CSF while Kaddurah-Daouk *et al.* (2011) study metabolites. The metabolite study found alterations in patients with AD for pathways related to the amino acids: tyrosine and tryptophan, purine, and vitamin E compound: tocopherol.

1.2.1.3 Cognitive tests

An obvious non-invasive biomarker is the use of cognitive testing. Measures such as the Mini-Mental State Exam (MMSE), Clinical Dementia Rating scale (CDR) and the ADAS-COG have been specially formulated to identify characteristics of people who have dementia (Folstein *et al.*, 1975; Hughes *et al.*, 1982; Rosen *et al.*, 1984). They are often split into functional categories allowing different symptoms to be identified and consequently a differential diagnosis to be made. For example, people with AD are more likely to show obvious deviations from age-associated decline in

episodic memory, semantic knowledge and some aspects of executive functions than people with other forms of dementia (Salmon & Bondi, 2010).

1.2.1.4 Limitations

Cognitive testing is not indicative of specific pathology and cannot detect pre-symptomatic changes. Additionally, collecting CSF measures and performing PET imaging is expensive and requires specialised facilities and staff. The Amyloid Imaging Task Force estimate that a single PET scan costs approximately $3,000 (Johnson *et al.*, 2013). Furthermore, internationally, there are mixed views on lumbar punctures with many patients experiencing anxiety and fearing pain. King & Rwegerera (2014) found that a majority of patients refused a lumbar puncture due to fear of pain and associated the procedure with paralysis and death. They concluded that attitudes were improved with education on the purpose, benefits and risks of the procedure. Meanwhile, Menéndez-González (2014) discuss that these preconceptions are culturally dependent. In Scandinavia lumbar punctures are viewed as routine procedures while in North America they are considered serious.

It is these reasons that motivate the research detailed in this thesis: Can we find a cost-effective biomarker of Aβ pathology that is more readily available than CSF and PET imaging? Ideally, a new biomarker could replace the use of PET scanning or CSF sampling. However, performance metrics of the test would need to be very high. For example, a test consistently achieving a sensitivity and specificity of >80% could be considered clinically useful in that its performance would be comparable to current markers in CSF (Jellinger *et al.*, 2008). A second possibility would be to use a new biomarker as a filtering step before performing confirmatory PET scans or lumbar punctures on those who are deemed at high risk for disease. Such a test would require slightly lower performance metrics, for example >70% sensitivity and specificity, and would reduce the number of people subject to the unpleasant and costly procedures of PET scanning and lumbar punctures.

1.2.2 Genetic Biomarkers

1.2.2.1 Genome Wide Association Studies

Genome Wide Association Studies (GWAS) test associations between SNPs and disease aiming to explain disease status using genetic variation. Such studies have rapidly increased the amount of data available on common genetic variants. Collections such as the National Human Genome Research Institute (NHGRI) Catalog aim to collate data from these studies (Welter *et al.*, 2014). In April 2016, the catalog contained 2,423 studies detailing 16,617 unique SNP-trait associations. A GWAS study in a simple case/control setting comprises multiple univariate logistic regression analyses, with an association deemed significant if a p-value significance threshold is reached. Conventionally, genome-wide significance is achieved for a p-value less than 5×10^{-8} although this varies between studies. GWAS in AD are beginning to have enough power to identify individual risk SNPs with Lambert *et al.* (2013) providing the most comprehensive GWAS to date identifying 20 risk loci. However, it is estimated that currently identified SNPs can only explain 16-33% of phenotypic variation (Lee *et al.*, 2013; Ridge *et al.*, 2013, 2016).

GWAS of Aβ and other pathological endpoints provide greater statistical power by reducing the heterogeneity contributed by controls who may have underlying pathology in a case/control analysis. Such studies are also highlighting promising results but sample sizes are considerably smaller so detailed validation is needed (Kim *et al.*, 2011; Cruchaga *et al.*, 2013). Kim *et al.* (2011) found four genetic loci that passed genome-wide significance testing for association with at least one CSF biomarker (Aβ, total tau or phosphorylated tau). In the larger study from Cruchaga *et al.* (2013) three genetic loci were significantly associated with tau.

1.2.2.2 Missing heritability

As genetic analysis becomes cheaper and consequently more readily available we can begin to investigate the coverage of common SNP chips. Heritability studies of AD and Aβ show that large amounts of genetic predisposition remain unexplained. This concept is known as missing heritability and is common across the majority of

complex diseases. There are several suggestions as to what is causing it. The main hypothesis is that coverage of commonly available SNP chips is sub-optimal. The method of next generation whole genome sequencing (WGS) begins to address this by enabling us to gather information on rare variants that was previously unavailable.

1.2.2.3 Whole Genome Sequencing

Next Generation WGS is a modern genetic technique that determines the complete DNA sequence of an organisms genome. In 2014, Illumina was the first company to offer this service for an estimated $1,000 per sample (Sheridan, 2014). In AD research, the first WGS data was made available in 2013 by the Alzheimer's Disease Sequencing Project (ADSP) (www.niagads.org). Projects such as 100,000 Genomes England (www.genomicsengland.co.uk) and the Longevity project are making use of this method to collect genetic data, alongside clinical and phenotypic information, for hundreds of thousands of individuals.

Cohort studies such as the Alzheimer's Disease Neuroimaging Initiative (ADNI) have been quick to get their samples sequenced. Nho *et al.* (2013) have used this data to identify a variant protective of hippocampal atrophy and to discover functional variants associated with changes in hippocampal volume, in participants with MCI (Nho *et al.*, 2013, 2015). Due to the novelty of WGS, studies emphasise the importance of replication in larger cohorts and longitudinal follow-up. Furthermore, WGS (among other sequencing methods) was used to discover that a T allele of the Triggering Receptor Expressed On Myeloid Cells 2 (*TREM2*) gene (SNP rs75932628) is associated with AD. It has been estimated based on Icelandic data that the risk variant occurs in only 0.63% of the population, making it extremely rare. It's effect size is second only to *APOE* (*TREM2* odds ratio = 2.92) (Jonsson *et al.*, 2013; Guerreiro *et al.*, 2013).

New SNP chips named exome chips have been designed to include variants identified through WGS and exome sequencing. In AD research, Sims *et al.* (2015) are beginning work on these chips to explore the idea that missing heritability is likely

to be explained by rare variants.

1.2.2.4 Polygenic Risk Score

There are three main methods investigating polygenic effects: the polygenic risk score, linear mixed effect regression models and linkage disequilibrium (LD) score regression (Purcell *et al.*, 2009; Yang *et al.*, 2010; Bulik-Sullivan *et al.*, 2015). The polygenic risk score (PGRS) was created to provide an approximate metric representing the cumulative effect of lower impact risk variants. Risk scores are generated per person using summary statistics from GWAS by multiplying per risk allele effect sizes by the number of risk alleles present. This is performed for all SNPs with a GWAS p-value below a pre-defined threshold. The summation of scores for each such SNP is used as the final PGRS. The p-value threshold is arbitrary and hence tools such as PRSice have been developed to investigate the most suitable threshold for a given analysis (Euesden *et al.*, 2015).

PGRS have proved informative of AD diagnosis and associated phenotypes but are yet to be included in a multi-modal analysis with other biomarkers of AD (Sabuncu *et al.*, 2012; Sleegers *et al.*, 2015).

1.2.3 Blood Biomarkers

One approach for finding a biomarker of early AD, that is more readily available than imaging or CSF measurements, is to search for a marker in blood.

1.2.3.1 Why blood biomarkers?

As blood and CSF interact through the BBB, albeit in a highly regulated manner, it is possible that a signal reflecting AD pathology could also be detected in blood (Ballabh *et al.*, 2004). However, due to the distance between brain and blood, it is likely any signal will be noisier than that seen in the CSF.

Additionally, it is thought that AD may affect the structure of the BBB allowing abnormal substances, or unusual quantities of compounds, to pass through (Marques *et al.*, 2013). The break down of the BBB in AD is the subject of much discussion. It

is well known that ageing leads to a decrease in the bodys ability to supply the brain with factors needed for normal function and to eliminate toxins. This is exacerbated in AD. As the BBB is an active barrier, not just a physical block, any problems are likely to be due to dysfunction rather than a complete barrier break down (Zhao *et al.*, 2015). On the other hand, Bien-ly *et al.* (2015) suggest that experimental drugs are not getting into the brain due to a lack of damage to the BBB. They explain that this lack of BBB permeability has been shown in several mouse models of AD and they have observed similar numbers of BBB breaches between human cases and controls.

Furthermore, the lymphatic system provides a route for signal originating in the brain to be seen in the blood. The lymphatic system is comprised of a series of vessels and nodes throughout the body with a primary function to transport lymph, a fluid containing white blood cells. In doing so, the lymphatic system helps the body to fight infection and has the secondary function of removing toxins and waste. More simply, the lymphatic system can be viewed as a drainage system for waste and toxins to be removed from organs and to exit the body via the blood. In 2014 it was discovered that the lymphatic system is present within the central nervous system of humans (Louveau *et al.*, 2015). Consequently, it is possible that a process occurring in the brain, such as build up of amyloid pathology, could be visible as signal in the blood.

The use of a blood test in disease diagnosis and monitoring is common. For example, blood glucose measurements have long been used to detect and monitor diabetes. However, the search for novel blood biomarkers from high-dimensional *omics* data has produced very few clinically usable tests, despite interest across a wide range of clinical areas. In oncology, the Oncotype DX test is a genomic test used to predict an individuals risk of breast cancer recurrence and response to treatment (www.oncotypedx.com). Although current evidence supports the conclusion that the Oncotype DX test is clinically valid, its impact on clinical outcomes is still being investigated (EGAPP Working Group, 2016). In other areas, *omics* discovery studies have been less successful. For example, Hoefer *et al.* (2015) explore the use

of *omics* biomarkers in atherosclerosis while Whiteley *et al.* (2012) review their use in predicting stroke. Both studies conclude that the clinical use of *omics* markers will depend on the reproducibility of their predictive ability across studies; a concept that has rarely been demonstrated. The use of blood based biomarkers has shown promise in research and limited clinical settings, and with improved methodologies and assays blood biomarkers may well reach clinical utility in the future.

1.2.3.2 Gene expression and pathways

Gene expression is the production of mRNA from a gene. mRNA are the family of RNA molecules that specify the conversion of genetic code stored on DNA to the final gene product, usually a protein, at the ribosome. Gene expression is usually quantified as the amount of corresponding mRNA molecules. By regulating gene expression a cell can control how much, or how little, of a protein it creates. Gene expression can be easily quantified in blood using microarray technology and has consequently been readily explored as a possible biomarker of AD. Han *et al.* (2013) state that the blood transcriptome is vital in the disease mechanism of AD. Their review also highlights the limited replicability at the single marker level with few transcripts identified as significantly and consistently altered across studies.

Genes can be grouped in to sets according to their biological functions; these sets can be described as pathways. It is common in genetic and gene expression analysis to perform a post-hoc pathway analysis to determine which pathways any differentially expressed gene belongs to. Han *et al.* (2013) report greater concordance between studies at the pathway level than at the the single gene level. In studies of AD and related phenotypes the most common differentially regulated pathways include mitochondrial dysfunction and immune activation (Lunnon *et al.*, 2012).

1.2.3.3 Proteins

Proteomic analysis is arguably the most well-researched area in the search for a blood biomarker of AD. This is largely because Aβ and tau are proteins themselves making blood proteins sensible biomarker candidates. Furthermore, as mentioned in section 1.2.1.2, the correlations of Aβ and tau between the brain and CSF are high.

Unfortunately, so far, the correlations observed when studying the blood are much weaker. This could imply that this disease signal cannot successfully pass through the BBB or secondly, that we are yet to design an assay sensitive enough to pick up small, AD related changes in molecular concentrations. Furthermore, amyloid is produced in the bone marrow and can be deposited in any organ or tissue. If an amyloid signal could pass through a damaged BBB it would be essential to ensure relevant signal, originating in the brain, was being detected and not confused by noise from other sources of amyloid. Consequently, the search began for other proteins in the blood that may associate with case/control status or disease characteristics such as cognitive test results, *APOE* genotype and, of particular interest, Aβ burden in the brain (Kiddle *et al.*, 2014).

At the beginning of this project (2013) four studies had been published investigating proteins associating with Aβ burden in four different cohorts (Thambisetty *et al.*, 2012; Kiddle *et al.*, 2012; Burnham *et al.*, 2013; Ashton *et al.*, 2015). Some proteins are significant in multiple analyses making them promising candidates for Aβ biomarkers; for example fibrinogen gamma chain (FGG) and pancreatic polypeptide (PPY) (Thambisetty *et al.*, 2012; Kiddle *et al.*, 2012; Burnham *et al.*, 2013; Ashton *et al.*, 2015). However, replication in larger cohorts, using a variety of technologies and biological samples is still necessary.

1.2.3.4 Metabolites

Metabolites are the products of any metabolic reaction and include substances such as vitamins, minerals, hormones and fats. Due to their small size it is thought that they are the most likely molecules to pass through the BBB, giving the possibility of a signal originating from the brain being measured in the blood. This effect is potentially magnified if the BBB is weakened by disease. Furthermore, the well-known risk gene for AD, *APOE*, codes for a protein involved in the transportation of cholesterol. This gives a second link to AD, as cholesterol is a metabolite, and highlights metabolites as potentially interesting molecules to study.

Metabolite studies have made several novel discoveries of possible markers for

AD. Ellis *et al.* (2015) highlight sphingolipid and glutamate metabolism as being altered in AD. They also highlight the difference in metabolism of molecules with antioxidant properties (including vitamin C and uric acid) in people with AD against controls. Several studies have used a case/control design to discover panels of metabolites associated with disease. For example, Proitsi *et al.* (2015) discovered a panel of 10 metabolites that differentiated between AD cases and controls. Further, Mapstone *et al.* (2014) studied early cognitive impairment in particular by using a similar 10 plasma metabolite panel to predict conversion from control status to amnestic MCI. Results from this study failed to replicate in serum samples from cohorts of a larger size (Casanova *et al.*, 2016). Despite interest in this area, no one has yet studied blood metabolite markers of Aβ .

1.2.3.5 Multi-modal biomarkers

There are many types of analyte within blood that can be measured and therefore have potential for use as a diagnostic tool (for example protein, metabolite and gene expression levels). If they are found to convey independent information on AD pathology, it is likely that a multi-modal biomarker model will be most accurate in identifying pre-symptomatic AD patients by offering improved predictive ability (Bazenet & Lovestone, 2012). However, it is important to study the individual marker types first, to identify the most promising markers, and to set a benchmark with which to compare a more costly multi-modal marker (Boulesteix & Hothorn, 2010). The literature to date contains several instances of demographic data (most commonly age and *APOE* ϵ4 status) being combined with cognitive scores and one type of blood biomarker. However, no multi-modal blood biomarkers of AD or related phenotypes have been researched. It has long been known that AD is a complex disease with many risk factors and therefore the suggestion of combining blood biomarkers is a sensible one. This thesis begins to address the lack of multi-modal blood biomarkers.

1.3 Conclusions

AD is a complex disease with many genetic and environmental risk factors. Ageing of the wordwide population means the number of people living with AD is sure to dramatically increase in coming years, presenting a large socio-economic problem. With no disease-modifying treatments available for AD it is essential to rethink drug discovery by making use of the 15-20 year period when AD pathology accumulates. As a starting point, AD clinical trial populations can be enriched for people with high Aβ burden and preventative trials can be performed in asymptomatic individuals with high Aβ burden. For this enrichment to be successful the relevant volunteers must be identified; a process that could be aided by the use of a blood based biomarker. The development of such a marker has been researched but clinical utility is yet to be achieved.

1.4 Thesis overview

This thesis begins to address the lack of a clinically useable blood biomarker of AD by investigating the issues of a lack of reproducibility and replicability. It also presents new methodologies to the field such as Bayesian techniques and multimodal approaches to search for new possibilities of blood-based biomarkers of AD. The cohorts used in this thesis are introduced in each chapter.

1.4.1 Single modality markers

As discussed above, research to discover a blood biomarker of AD is well-established with the majority of research groups focusing on single modality approaches. Results frequently differ between studies indicating that replication of results is lacking and should be the key aim of modern research (Bazenet & Lovestone, 2012). If replication is not possible, improved methodologies and novel study designs should be investigated.

1.4.1.1 Chapter 2: Blood protein markers of neocortical Aβ burden

The paper presented in Chapter 2 has been published in the Journal of Alzheimer's Disease (Voyle *et al.*, 2015). It begins to address the replication issues in proteomics biomarker research by taking a candidate approach to replicating previously identified blood protein markers of Aβ burden.

1.4.1.2 Chapter 3: A pathway based approach for analysing gene expression for AD diagnosis

Several gene expression candidates have appeared promising as biomarkers for AD. However, between studies there is minimal concordance in the markers identified (Han *et al.*, 2013). The paper in chapter 3 has been published in the Journal of Alzheimer's Disease (Voyle *et al.*, 2016a). It investigates the hypothesis that by creating pathway level scores of gene expression I may create a more robust, and hence reproducible, biomarker of AD case/control diagnosis.

1.4.2 Multiple modality markers

Although single modality biomarkers have shown some promising associations with diagnostic and endophenotypic endpoints of AD, nothing has reached clinical utility. In this thesis, I hypothesise that combining independent signal from several modalities of biomarker may provide potential utility. To test this hypothesis in AD, multi-modal cohort data has been used to model Aβ and tau burden.

1.4.2.1 Chapter 4: Blood metabolite markers of neocortical Aβ burden

To date, studies of metabolites as possible biomarkers for AD are based only on a diagnostic endpoint. Due to the heterogenous nature of AD this is often inappropriate, particularly as the pathology of disease can develop 15-20 years in advance of any symptomatic changes. It is therefore imperative that the methodologies used in these studies are extended to the investigation of an endophenotypic endpoint of AD. This issue is addressed in chapter 4; the first study to search for associations between Aβ burden in the brain and metabolite markers in the blood. This study

has been published in Translational Psychiatry (Voyle *et al.*, 2016b).

In chapter 4 I present a study that is looking to discover a panel of metabolites to predict Aβ burden in the brain. This paper also presents an enriched panel with the addition of two reasonably well-replicated proteins: FGG and PPY. On addition of FGG, accuracy in test data improves by 7% providing evidence that the idea of multi-modal research warrants further investigation.

1.4.2.2 Chapter 5: Do peripheral markers help to predict amyloid burden in a non-demented population? A Bayesian approach

The paper in chapter 5 focuses on using a PGRS to predict amyloid. A novel, Bayesian approach was used for this analysis to predict normal and abnormal amyloid as measured in CSF. The Bayesian analysis was used to compare models built using non-informative prior distributions for age, diagnosis and *APOE* to those built using informative priors based on results from a meta-analysis by Jansen *et al.* (2015). This analysis aims to meet the multi-modal objectives of this project by combining the PGRS with information from gene expression, protein and metabolite measurements in a subset of individuals as supplementary analysis. This work is in submission at PLoS Medicine.

1.4.2.3 Chapter 6: Genetic risk as a marker of amyloid-β and tau burden in cerebrospinal fluid

Chapter 6 presents work focused on predicting a more 'AD like' representation of pathology. I have used a PGRS to predict normal or abnormal amyloid, tau and a combination of the two pathologies. The latter analysis compares individuals with normal CSF amyloid and tau to those with abnormal CSF amyloid and tau. This chapter aims to meet the multi-modal objectives of this project by incorporating measures of tau in blood plasma in a sub-study. This work has been accepted for publication at the Journal of Alzheimer's Disease.

1.5 Details of methodologies

As this thesis is presented as a series of published or submitted papers, the methods sections are included within each chapter. In the methodology sections below I outline some of the more complex methods in more detail.

1.5.1 Statistical methodologies

While completing the work that contributed to this thesis I co-authored a book chapter for *Practical Psychiatric Epidemiology, Second Edition (in press)*. The chapter (Statistical Methods in Epidemiology: Bio-informatician's perspective) outlines basic data handling, exploration and analysis methods vital for psychiatric epidemiology research. A copy of the full chapter can be found in Appendix C and the following sections are exact replicas of some of the sections I authored (indicated with quotation marks).

1.5.1.1 Training, testing and validation

'A common problem of supervised learning methods is over-fitting. An algorithm can be trained for too long on one particular data set so that it fails to generalise the information learned to similar data sets. A number of methods are available for evaluating machine learning results and showing the results are general enough to be applied to other data (Hand *et al.*, 2001). To successfully train a supervised learning algorithm, one should aim to generate three data sets from the original data set.

- A training set: used to train the algorithm.
- A validation set: to track how well the algorithm is generalizing and to perform parameter tuning.
- A test set: an unseen data set on which the finalised algorithms performance is tested.

As machine learning algorithms require substantial training data, the usual distribution of the training, validation and testing sets is 2:1:1. However, in the case of

limited data availability, different cross-validation (CV) methods can be employed.'

1.5.1.2 Cross Validation (CV)

'In cross-validation the data is divided into a number (n) of 'folds'. Each fold is treated as the validation dataset in turn, with the remaining n-1 folds being used as training data. CV is especially useful for smaller datasets (Kohavi, 1995). The performance of the classifier on each fold is measured and then a final accuracy is calculated based upon the average of all n folds.'

1.5.1.3 Random Forests (RF)

'Random forest (RF) is a supervised classifier consisting of multiple decision trees (Breiman, 2001). The final class assigned to an observation is the modal class selected by the multiple decision trees. RF combines two machine learning methods: bootstrap sampling and random feature selection. Each tree is created from a bootstrap sample of the training data. OOB [out of bag] data is used to obtain an unbiased estimate of the error during the training. However, rather than using all features, RF randomly selects a subset of input variables to decide what decision should be made at each node of the tree. Advantages of random forest classifiers include the fact that the error can be balanced when the class population sizes are imbalanced and over-fitting can be avoided. There are also good methods available for handling missing data.'

1.5.1.4 Partial Least Squares (PLS)

'Partial least squares (PLS) modelling is very similar to PCA [Principal Components Analysis]. Where PCA looks for orthogonal hyper-planes that explain maximal variance in the predictors, PLS looks to explain maximal covariance between the predictors and the outcome (Wold, 2004). It is this reliance on the outcome measure that means PLS modelling is supervised and is particularly suited for prediction problems. PLS modelling is useful when the number of predictors is greater than the number of samples as it reduces the size of the predictor space by creating components.'

1.5.1.5 Support Vector Machines (SVMs)

'Support vector machines (SVMs) are a supervised learning classifier developed by Cortes & Vapnik (1995). They have been shown to be very accurate in many disciplines including bioinformatics, benefiting from the ability to handle high dimensional data with a small number of instances, finding a good balance between training set accuracy and test data error. For a given set of training vectors labelled with two classes, a SVM can find the optimal linear hyper-plane that maximizes the margin between the two classes. An example of SVM classification in two dimensions is given in Figure 1.3. SVMs can be extended to provide non-linear classification through the application of a kernel function and to multi-class classification but that is not discussed here.'

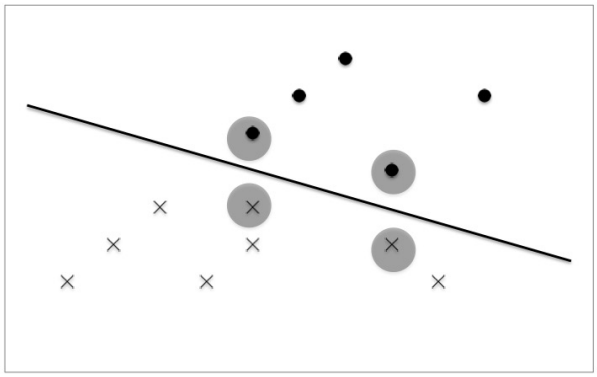


Figure 1.3: An example of a SVM classifying between two classes (circles and crosses). The points acting as support vectors are highlighted. Reproduced from *Practical Psychiatric Epidemiology, Second Edition.*

1.5.1.6 Bayesian Logistic Regression

Using a Bayesian methodology enables us to include historical information when model building. Bayesian statistics is based on the ability to update some prior beliefs with new data to form posterior beliefs. In many cases, the integrals needed to perform this calculation can only be estimated using numerical methods. The most common, and computationally efficient, way to do this is through Markov Chain Monte Carlo (MCMC) sampling. MCMC draws values from a suggested distribution forming a Markov Chain meaning each draw depends only on the previous one. In

this thesis we use a Metropolis sampling method where each value is accepted or rejected based on some acceptance criteria. It is important to maintain moderate levels of rejection so the Markov chain explores the posterior distribution. Over time, the set of accepted values will converge to a distribution that can be used for inference.

In the case of Bayesian logistic regression analysis, MCMC sampling is performed to estimate the effect size of each term in a model. The means of the MCMC samples, after convergence, can then be extracted and used as effect sizes for prediction in new data. In this thesis we investigated informative prior distributions, derived from historical data, and non-informative priors. Non-informative priors are designed to portray minimal information so often have large variances and can be considered 'flat'. In the studies presented in chapter 5 of this thesis we found that Bayesian modelling using non-informative priors yielded similar results to a frequentist logistic regression, as expected.

1.5.2 Scientific methodologies

Chapters 2 and 4 make use of novel quantification techniques for proteins and metabolites respectively: SOMAscan and LC-MS/MS. More information on these technologies is given below.

1.5.2.1 SOMAscan technology

The need for advances in proteomic technologies is common across many fields including pharmaceutical development, diagnostics and biomarker discovery. Modern proteomics technologies are generally limited by two main factors: lack of sensitivity and inability to achieve high-throughput. The SOMAscan platform (SomaLogic Inc, Boulder, CA) aims to address these limitations through the use of Slow Off-Rate Modified Aptamers (SOMAmers). SOMAmers are chosen to bind to discrete molecular targets with improved specificity and are currently being multiplexed at over 1000 SOMAmers per sample. A more in-depth idea of how the SOMAscan technology works is given in Figure 1.4 (Kraemer *et al.*, 2011). This details how

the original protein signal is converted to a nucleotide signal that can be measured using standard DNA quantification methods.

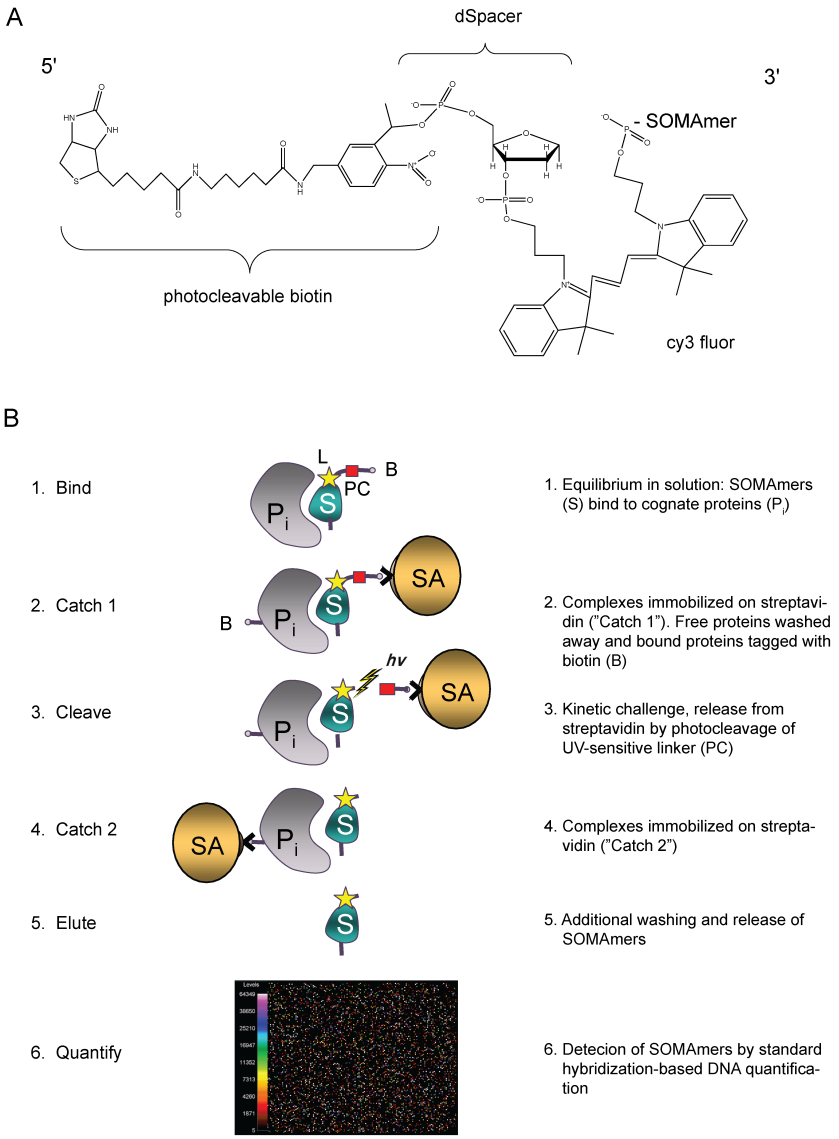


Figure 1.4: SOMAmer-based assay reagent and assay principles. Reproduced from Kraemer *et al.* (2011).

1.5.2.2 Liquid chromatography tandem mass spectrometry (LC-MS/MS)

LC-MS/MS combines the physical separation properties of liquid chromatography with the mass analysis capabilities of mass spectrometry. The method provides high sensitivity to detect a variety of components from a complex mixture such as serum or plasma. Tandem mass spectrometry comprises two mass analysis steps. The first step investigates mass of the precursor ion while the second step provides

fragmentation patterns for the product ions. It is this dual mass analysis that provides improved sensitivity over MS or LC-MS as well as improved structural information from the fragmentation patterns. The LC-MS/MS method used in this thesis has been previously published and can be referred to for more information (Whiley *et al.*, 2014).

Chapter 2

Blood protein markers of neocortical amyloid-β burden: A candidate study using SOMAscan technology

Journal of Alzheimer's Disease 46 (2015) 947–961
DOI 10.3233/JAD-150020
IOS Press

Blood Protein Markers of Neocortical Amyloid-β Burden: A Candidate Study Using SOMAscan Technology

Nicola Voyle[a,b,*], David Baker[c], Samantha C. Burnham[d], Antonia Covin[c], Zhanpan Zhang[e], Dipen P. Sangurdekar[e], Cristina A. Tan Hehir[e], Chantal Bazenet[a,f], Simon Lovestone[a,g], Steven Kiddle[a,b,1], Richard J.B. Dobson[a,f,*,1] and the AIBL research group[h]
[a]*Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK*
[b]*MRC Social, Genetic and Developmental Psychiatry Centre, King's College London, London, UK*
[c]*Janssen R&D, Neurosciences, Titusville, NJ, USA*
[d]*CSIRO Digital Productivity and Food and Nutrition Flagships: eHealth, Floreat, WA, Australia*
[e]*GE Global Research, Niskayuna, NY, USA*
[f]*NIHR Biomedical Research Centre for Mental Health and Biomedical Research Unit for Dementia at South London and Maudsley NHS Foundation, London, UK*
[g]*Department of Psychiatry, Oxford University, Oxford, UK*
[h]*http://www.AIBL.csiro.au*

Handling Associate Editor: Gary Arendash

Accepted 26 March 2015

Abstract.

Background: Four previously reported studies have tested for association of blood proteins with neocortical amyloid-β burden (NAB). If shown to be robust, these proteins could have utility as a blood test for enrichment in clinical trials of Alzheimer's disease (AD) therapeutics.
Objective: This study aimed to investigate whether previously identified blood proteins also show evidence for association with NAB in serum samples from the Australian Imaging, Biomarker and Lifestyle Flagship Study of Ageing (AIBL). The study considers candidate proteins seen in cohorts other than AIBL and candidates previously discovered in the AIBL cohort.
Methods: Our study used the SOMAscan platform for protein quantification in blood serum. Linear and logistic regressions were used to model continuous NAB and dichotomized NAB respectively using single proteins as a predictor. Multiple protein models were built using stepwise regression techniques and support vectors machines. Age and *APOE* ε4 carriage were used as covariates for all analysis.
Results: Of the 41 proteins previously reported, 15 AIBL candidates and 20 non-AIBL candidates were available for testing. Of these candidates, pancreatic polypeptide (PPY) and IgM showed a significant association with NAB. Notably, IgM was found to associate with continuous NAB across cognitively normal control subjects.

[1]These authors are considered joint senior authors.

*Correspondence to: Nicola Voyle and Richard J.B. Dobson, Social, Genetic and Developmental Psychiatry Centre, Institute of Psychology, Psychiatry and Neuroscience, King's College London, De Crespigny Park, Denmark Hill, London SE5 8AF, UK. Tel.: +020 7848 0924; Fax: +020 7848 0866; E-mails: nicola.voyle@kcl.ac.uk (Nicola Voyle) richard.j.dobson@kcl.ac.uk (Richard J.B. Dobson).

ISSN 1387-2877/15/$35.00

Conclusions: We have further demonstrated the association of PPY and IgM with NAB, despite technical differences between studies. There are several reasons for a lack of significance for the other candidates including platform differences and the use of serum rather than plasma samples. To investigate the possibility of technical differences causing lack of replication, further studies are required.

Keywords: Alzheimer's disease, amyloid plaques, blood, positron emission tomography scan, proteomics

INTRODUCTION

The pathology of AD is characterized by three features in the brain: the aggregation of amyloid-β (Aβ) into plaques, the presence of hyperphosphorylated tau in the form of tangles, and the occurrence of neuron loss leading to brain atrophy [1, 2]. There is debate as to how these processes interact to cause symptomatic AD. A popular theory is the amyloid cascade hypothesis: that Aβ deposition is central to disease development [3]. Studies show that Aβ plaques begin to develop up to 20 years prior to clinical diagnosis with their growth reaching a plateau as clinical symptoms arise [4, 5]. The presence of AD pathology can be investigated through characteristics that are informative of AD diagnosis; endophenotypes. Measurements of analytes in cerebrospinal fluid (CSF) and molecular imaging by positron emission tomography (PET) scans are examples of endophenotypes that are biomarkers for AD. The disadvantages of the procedures involved in attaining these measurements include their invasive or expensive nature and that they require specialized administration.

Existing treatments for AD provide short-term symptomatic relief, in a subset of patients, and trials of potential disease modifying treatments are failing [6]. Two suggestions have been made to address this shortcoming: the use of Aβ as a companion diagnostic in tertiary prevention trials (those investigating disease modifying or symptomatic treatments) and in secondary prevention trials (investigating treatments to prevent the disease from occurring in the first place). Firstly, we consider the use of a companion diagnostic in a normal clinical trial. To date, trials of Aβ targeting drugs have recruited patients with a clinical AD or mild cognitive impairment (MCI) diagnosis and unknown brain Aβ burden [7–10]. As Aβ PET scans became more readily available, toward the end of such trials, Aβ burden was measured and it was discovered that some of these AD patients had low Aβ burden (the target pathology). For example, in a trial of Bapineuzumab, a humanized anti-Aβ monoclonal antibody, approximately 14% of subjects had low Aβ burden [9]. Consequently, it has been suggested that Aβ PET or CSF measures are used as a companion diagnostic, with elevated brain Aβ becoming an eligibility requirement for Aβ-lowering trials. A panel of blood biomarkers could enable cost-effective enrichment and identification of trial participants with Aβ pathology for subsequent confirmatory Aβ CSF tests or PET scans.

A second use of a blood-based biomarker would be in recruitment for secondary prevention trials. It is hypothesized that the brains of people recruited with a diagnosis of AD or MCI are often too damaged for medication to have an effect. Treatments could have a higher chance of success in subjects with no clinical symptoms, under the assumption that it is easier to delay, rather than reverse, the development of AD pathology. Aisen et al. expand on this idea stating that there has been no medical advancements in the last decade in terms of treating AD [11]. They suggest basing recruitment of patients into clinical trials on the presence of biomarker defined pre-symptomatic AD; three such trials had begun by the end of 2014. As above, using a blood biomarker as an enrichment filter for PET scans and CSF markers could increase the efficiency of clinical trial recruitment by reducing screen failures. A simple blood test has the greatest potential impact in streamlining secondary prevention studies as screen failure rates due to low Aβ are expected to be higher ($\sim$78.1% versus $\sim$12%) in asymptomatic individuals than in subjects with AD (Ossenkoppele et al. and Jansen et al., unpublished results).

Consequently, there is high demand for a blood-based biomarker of AD pathology [12]. It has been shown that measures of Aβ in the brain are highly negatively correlated with those in the CSF [13]. However, overall, studies of blood Aβ do not show similar correlations [14]. It is therefore necessary to search for other analytes within the blood that may be associated with Aβ in the brain. Recent studies provide some evidence that proteins in blood show such a link to other AD endophenotypes and to Aβ burden in the brain [15–19]. However, there is a need to investigate these results further in independent datasets of a larger size, using different technologies and different biological samples. We begin to address this here.

This study aimed to further test findings in the literature of proteins in blood plasma associating with brain Aβ burden [15–17, 19]. We used a radioactive marker of Aβ, [^{11}C]-Pittsburgh compound B (PiB), combined with PET scanning to quantify the amount of brain Aβ burden and the SOMAscan proteomics platform [20] for quantification of protein in blood samples.

With over 1,000 protein analytes, the SOMAscan proteomics platform has been used in high-throughput biomarker discovery studies. For example, in AD research it has been used by Sattlecker et al. and Zhao et al. to discover potential plasma biomarkers of AD diagnosis, MRI measures, and/or rate of cognitive decline [18, 21]. In this study we focused our approach on candidate-based analysis, selecting only proteins previously found to associate with NAB. This was done to ensure that there was sufficient statistical power to detect weaker signal that could have been missed had the entire SOMAscan panel been used. This analysis was repeated within just the cognitively normal control subjects to evaluate whether this test would have utility to enrich for Aβ-positivity in secondary prevention trials.

MATERIALS AND METHODS

Cohort

AIBL is a prospective, longitudinal study of subjects aged over 60 years. The dual center study recruits subjects with an AD diagnosis as well as both control subjects and those with MCI with the aim of identifying factors that lead to subsequent AD development. Details of study design and enrolment are given elsewhere [17]. The present study comprised a subset of 198 subjects from AIBL, enriched for controls.

Proteomics

Blood draws (80 ml) were taken after overnight fasting and serum was collected in Sarsedt s-monovette tubes. The samples were centrifuged at 1,800 g for 15 min at 20°C and immediately frozen at −80°C. They were stored (long-term) in liquid nitrogen [22]. Protein levels in the blood serum were analyzed using the SOMAscan platform (SomaLogic Inc, Boulder, CA). The methods used in this assay are outlined in detail by Kraemer et al. but, in brief, proteins were measured using Slow Off-rate Modified Aptamer (SOMAmer)-based capture arrays using a sample of less than 10 μL per run [23]. SOMAmers are nucleotides that have been chemically modified to address two issues: hard to capture proteins and non-specific binding. Firstly, the nucleotides are given protein-like properties resulting in high affinity for hard to capture proteins. Secondly, aptamers with slow dissociation rates are used allowing disruption of non-specific binding. These nucleotides are used to transform a protein signal to a nucleotide signal that can be quantified using relative florescence on microarrays. For full quantification details of this assay, see Gold et al. [20]. This study used SOMAscan Version 2 which captured information on 1,001 proteins and protein complexes.

Quality control was performed at the sample and SOMAmer level, and involved the use of control SOMAmers on the microarray and calibration samples. Hybridization controls measured sample-by-sample variation in hybridization while the median signal over all SOMAmers measured technical variability. Scale factors of these two metrics were used to normalize across all samples with acceptance criteria of 0.4 to 2.5 based on historic trends. SOMAmer by SOMAmer calibration occurred through the repeated measurement of calibration samples. Historic values were used to generate a calibration scale factor the acceptance criterion for which was: 95% of SOMAmers must have a calibration scale factor within 0.4 of the median.

Imaging

The measures of NAB used in this study were collected by PiB PET scans. The PiB imaging methodology is detailed elsewhere [24]. Each scan was spatially normalized by Rowe et al. to a customized PiB-PET template in the Montreal Neurological Institute reference space using Statistical Parametric Mapping 8 (SPM8; Wellcome Trust Centre for Neuroimaging, London, UK). Standardized uptake value ratios (SUVRs) were then created by computing the ratio of PiB retention in the whole brain to that in the grey matter [25].

Identification of candidate blood protein markers of brain Aβ burden

We identified candidate proteins through a search of the literature for studies investigating the relationship between blood proteomics and Aβ burden in the brain. Any proteins identified as significant in the study conclusions were selected and matched to the AIBL SOMAscan data using the UniProt ID.

Statistical analysis

All statistical analysis was performed in R (version 3.1.1) [26].

Differences in proteomic signal between platforms and sample types

As the candidates were selected based on their association with NAB in plasma, we aimed to confirm that serum is an appropriate surrogate by comparing proteomic signal from the SOMAscan platform between blood plasma and serum samples using age as an outcome. We also studied differences between the SOMAscan platform and Myriad's Rules Based 415 Medicine Multi-analytes Profile (RBM MAP), one of the discovery proteomics platforms. To do so, we used proteomic data generated using the RBM MAP platform from the original publication of a blood biomarker in the AIBL study; methods described by Burnham et al. [17].

Menni et al. have studied associations of age with protein levels from plasma samples using the TwinsUK cohort and SOMAscan version 3 [27]. Running random intercept linear regression, adjusting for family relatedness, identified proteins associated with chronological age. We performed a similar analysis using standard linear regression on our serum samples from the AIBL cohort. Within each cohort (TwinsUK and AIBL) the proteins were ranked by p-value from the regressions and the R package OrderedList was used to compare the rank of proteins present on both versions of the platform [28].

To assess the concordance of proteomic signal between SOMAscan and RBM MAP, we implemented an identical method. OrderedList was used to compare lists of proteins ranked according to their associations with age and gender.

Analysis overview

An overview of the analysis is given in Fig. 1. Proteins discovered in studies other than AIBL [15, 16] were analyzed separately to proteins discovered in AIBL using different biological samples (serum instead of plasma) [17, 19]. This split was implemented as the association of a non-AIBL candidate with NAB in AIBL would provide a fully independent replication of those findings, whereas the association of an AIBL candidate with NAB would only show that the marker could also be measured using SOMAmers on serum samples. Each analysis tested the protein levels against the continuous endpoint of SUVR as well as a dichotomized high/low NAB endpoint. A cut off of 1.3 was used to distinguish between high and low NAB [17]. Use of a continuous endpoint gives increased power in statistical analysis in comparison to a categorical measure [29]. Additionally, each analysis studied all samples and control samples alone.

Subjects were randomly split into a training (66%) and test set (34%) using the createDataPartition function in the R package caret [30]. The split ensured that a range of values of NAB were present in each set for the continuous analysis by splitting the subjects into groups based on NAB percentiles and sampling within these groups. For the dichotomized analysis, the split ensured that the proportion of high and low NAB was similar between the two sets by sampling within the high and low groups.

Data pre-processing

Data pre-processing was performed separately for each analysis. In both cases, all protein data was subject to a natural logarithm transformation so each protein followed an approximate normal distribution. Per sample, per protein outliers were then identified as values lying outside of 6 standard deviations (SD) of the mean protein value and were set to missing. Any sample with more than 20% missing data was removed. Missing data in the train and test sets was imputed separately using 10 nearest neighbors.

In some cases, there were multiple SOMAscan probes available for one UniProt ID. In this instance, a principal components analysis (PCA) was performed across these probes and only the first principal component (PC1) was included in the analysis. This was performed on all samples and on control samples alone, and separately for train and test sets. Additionally, the colinearity of such probe sets was investigated using Pearson's correlation coefficients.

Single protein analysis

Models were built using the training set. The continuous and dichotomized NAB endpoints were regressed against age, APOE $\varepsilon 4$ load and each protein in turn using linear and logistic regression, respectively. Age and APOE $\varepsilon 4$ load, defined as the number of $\varepsilon 4$ alleles in a subjects genotype, were included as they were statistically significant in this population (see Table 2).

In all cases, Benjamini Hochberg corrected p-values from the full model fit (q-values) were tested against a significance level of 0.1, to account for false discov-

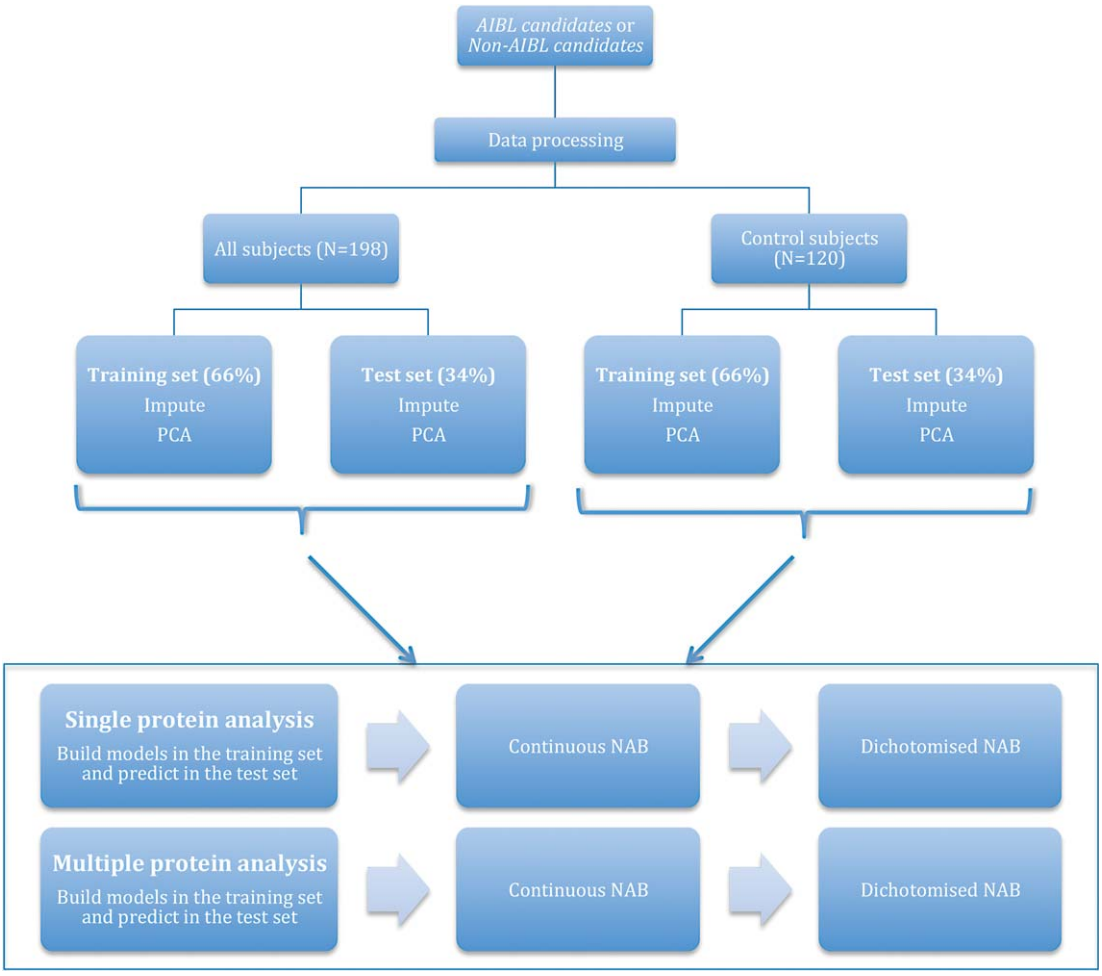


Fig. 1. Overview of data analysis.

ery. Although the hypothesis tests performed were not independent, the Benjamini Hochberg correction gave an acceptable approximation of the false discovery rate [31].

Model fit was assessed by exploring the residuals of each model to ensure normality, constant variance, zero expectation, and independence. For logistic regression models, deviance residuals were used.

Each model was used to predict NAB in the test set and model statistics were calculated. These were R^2 and root mean squared error (RMSE) for linear regression and accuracy, sensitivity and specificity for logistic regression. For each model, statistic permutation testing was run with 1,000 permutations of the NAB values, to create a non-parametric, empirical *p*-value. In each permutation, the relationship between NAB endpoint, age, and APOE $\varepsilon 4$ load was maintained. Benjamini Hochberg corrected *p*-values (q-values) were tested against a significance level of 0.1.

Multiple protein analysis

The primary multiple protein analysis was parametric. Models were built using the training set. As in the single protein analysis, continuous and dichotomized NAB endpoints were regressed against age, APOE $\varepsilon 4$ load, and all proteins together (15 non-AIBL candidates and 20 AIBL candidates) using linear and logistic regression respectively. The models were then simplified using stepwise regression based on Akaike's Information Criteria (AIC). Simplified linear regression models were examined for variable importance using the Lindemann, Merenda, and Gold (LMG) measure [32]. This metric represents the R^2 of an individual regressor by averaging over all orders of regressors. Each model was used to predict on the test set and create model statistics with associated empirical *p*-values, as detailed previously. Model fit was assessed in the same manner as for the single protein analysis.

As a secondary multiple protein analysis we investigated the possibility of non-linear associations and interactions by performing a non-parametric multiple protein analysis using support vector machines (SVMs) in the R package kernlab [33]. We used an SVM with radial kernel to build models in the training set. The train function from the R package caret was used to estimate model parameters: The kernel parameter (sigma) was directly estimated and regularization parameter (C) was tuned using 25 bootstraps. RMSE was used as the optimization metric for continuous NAB while kappa was used for the dichotomized endpoint. Each model was used to predict on the test set.

RESULTS

Cohort demographics

For this study, we used SOMAscan data generated on serum samples from the AIBL cohort. Table 2 summarizes the demographics of this population. As expected, characteristic features of AD such as Mini-Mental State Examination score and Clinical Dementia Rating were significantly associated with continuous NAB and significantly different between high and low NAB groups. Age and APOE ε4 load (defined as the number of APOE ε4 alleles in a subject's genotype) were also significant and hence accounted for in all analyses.

Overall, demographics in the control population mirrored those across all samples and hence the same terms were covaried for in the control only analyses (see Supplementary Table 1). As expected, measures of APOE ε4 were not significant in this sub-population.

Differences in proteomic signal between platforms and sample types

We began by comparing proteomic signal associated with age in plasma samples from the TwinsUK cohort and serum samples from AIBL. 975 SOMAscan probes matched between the Twins UK (Version 3, 1,129 SOMAmers) and AIBL (Version 2, 1,001 SOMAmers) datasets mapped using SOMAscan ID. Comparison of the protein lists ranked by association with age gave a significance of similarity p-value of <0.001. The analyses were not adjusted for any covariates.

We also made a comparison between the SOMAscan platform and RBM MAP platform using samples from the AIBL cohort. The RBM MAP data contained 151 proteins, 119 of which could be mapped to a UniProt ID. All 1,001 proteins on the SOMAscan panel were mapped to a UniProt ID. There were 88 proteins that overlapped between the two datasets, mapped by UniProt ID. Comparison of the protein lists ranked by association with age gave a significance of similarity p-value of 0.162. An identical analysis with proteins ranked by association with gender gave a significance of similarity p-value of 0.201.

Identification of candidate blood protein markers of brain Aβ burden

We identified four previous studies investigating a proteomic signal in the blood associated with Aβ burden in the brain. Previously reported studies used blood plasma whereas this study uses blood serum samples. Thambisetty et al. discovered six proteins associated with brain Aβ burden in a group of non-demented, older individuals [15]. Kiddle et al. used data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) in a discovery analysis to find 16 proteins associated with brain Aβ burden [16]. Significant proteins from these two papers that were also present in the AIBL SOMAscan data [15] are termed non-AIBL candidates and are given in Table 1.

Additionally, two discovery analyses have already been performed on AIBL subjects using different protein measurement technologies; RBM MAP [17] and mass spectrometry [19]. Respectively, these studies highlighted 8 and 17 proteins (a prioritized set) as being significantly associated with NAB. The proteins that were also present in the AIBL SOMAscan data [20] are termed AIBL candidates and are given in Table 1. See Supplementary Figure 1 for an overview of protein numbers.

In some cases, there were multiple SOMAscan aptamers for one UniProt ID: C3 (P01024) had six matches in the AIBL SOMAscan dataset, APOE (P02649) had three, and Fibrinogen ([P02671, P02675, P02679]) had two. All are accounted for in this analysis using PCA as detailed previously. Eleven of the 41 proteins were not present on the SOMAscan panel. AβPP was included as an additional protein of interest as $A\beta_{1-42}$ is not present on the SOMAscan panel.

Data pre-processing

There was no observed pattern in the missing data nor in the demographics of the subjects with data set to missing so it was assumed the data were missing at random. No subjects were removed during data pre-processing.

The protein data was generated in one batch and the PET scans were performed using one model of

Table 1
Candidate proteins

Protein	Gene	UniProt ID	Ashton et al. [19]	Burnham et al. [25]	Kiddle et al. [24]	Thambisetty et al. [23]	Total	Comments
Complement C3	C3	P01024	X		X	X	3	5 probes present.
Apolipoprotein E	APOE	P02649			X	X	2	3 probes present.
Fibrinogen (α, β, γ)	FG	P02671, P02675, P02679	X		X		2	3 probes present. Separate signal from α and γ chains [19]
Haptoglobin	HP	P00738	X			X	2	
Pancreatic Polypeptide	PPY	P01298		X	X		2	
$A\beta_{1\text{-}42}$				X			1	Not present in AIBL SOMAscan.
Albumin	ALB	P02768				X	1	
Alpha 1 Antitrypsin	SERPINA1	P01009			X		1	
Alpha 2 Macroglobin	A2M	P01023	X				1	
Apolipoprotein A	LPA	P08519	X				1	Not present in AIBL SOMAscan.
Apolipoprotein A1	APOA1	P02647	X				1	
Apolipoprotein A4	APOA4	P06727	X				1	Not present in AIBL SOMAscan.
Apolipoprotein L1	APOL1	O14791	X				1	Not present in AIBL SOMAscan.
AXL Receptor Tryosine Kinase	AXL	P30530			X		1	Not present in AIBL SOMAscan.
C-Peptide	INS	P01308			X		1	Not present in AIBL SOMAscan.
C4b Binding Protein Alpha Chain	C4BPA	P0C0L4	X				1	
Chemokine Ligand 13	CXCL13	O43927		X			1	
Clusterin	CLU	P10909	X				1	
Complement Factor B	CFB	P00751	X				1	
Complement Factor H	CFH	P08603	X				1	
Complement Factor H Related Protein 1	CFHR1	Q03591	X				1	Not present in AIBL SOMAscan.
Cortisol	(N/A)	(N/A)			X		1	Not present in AIBL SOMAscan.
Free Thyroxine	(N/A)	(N/A)		X			1	Not present in AIBL SOMAscan.
Gelsolin	GSN	P06396	X				1	
Hemopexin	HPX	P02790	X				1	
Histidine-rich Glycoprotein	HRG	P04196	X				1	
Ig Gamma 1 Chain C Region	IGHG1	P01857				X	1	
Immunoglobulin E	(N/A)	(N/A)			X		1	Not present in AIBL SOMAscan.
Immunoglobin M 1	IGHM	P01871		X			1	
Interleukin 13	IL13	P35225			X		1	
Interleukin 17	IL17	Q16552		X			1	
Interleukin 3	IL3	P08700			X		1	
Leptin	LEP	P41159			X		1	
Macrophage Inflammatory Protein (1α)	CCL3	P10147		X			1	
Matrix Metalloproteinase 9 Total	MMP9	P14780			X		1	
Plasminogen	PLG	P00747				X	1	
Serotransferrin	TF	P02787	X				1	
Serum Amyloid p-Component	APCS	P02743			X		1	
Vascular Cell Adhesion Protein	VCAM1	P19320		X			1	
Vitronectin	VTN	P04004			X		1	Not present in AIBL SOMAscan.
Von Willebrand Factor	VWF	P04275			X		1	
Amyloid-β Protein Precursor	AβPP	P05067					0	Included as an alternative to $A\beta_{1\text{-}42}$

Table 2
Population demographics

	Total ***n*** = 198	**Overall** ***p***-value	**High NAB** ***n*** = 107	**Low NAB** ***n*** = 91	**High/Low** ***p***-value
Gender (% female)	50	0.96	49	52	0.78
APOE status (% of *APOE* ε4 positive)	51	<0.001	61	50	0.004
APOE ε4 load (% with loads 0; 1; 2)	49; 43; 8.1	<0.001	39; 48; 13	60; 37; 2.2	0.002
Median age [IQR] (years)	72 [14]	<0.001	76 [11]	68 [12]	<0.001
Median MMSE score [IQR]	28 [3.0]	<0.001	28 [4.0]	29 [2.0]	<0.001
Global CDR status (%>0)	44	<0.001	64	21	<0.001
Clinical diagnosis (% with diagnosis AD; MCI; Control)	14; 26; 61	0.28	25; 35; 40	0; 15; 85	<0.001

Individuals were positive for APOE ε4 if at least one APOE ε4 allele was seen in their genotype.
APOE ε4 load was the number of ε4 alleles seen in a subject's genotype.
IQR, inter-quartile range; MMSE, Mini-Mental State Examination; CDR, Clinical Dementia Rating.
Overall *p*-value: Result of the Kendall tau test for dependence between SUVR and the demographics variable.
High/low *p*-value: Kruskal Wallis Chi-Squared was used to test between high and low groups for continuous data.
High/low *p*-value: Fishers exact was used to test between high and low groups for categorical data.

Table 3
Candidate proteins of the same UniProt ID

UniProt ID	**Proteins**	**Minimum-Maximum Correlation**
P01024	C3; C3adesArg; C3b; iC3b; C3a; C3d	0.468 – 0.964
P02671	Fibrinogen; D-dimer	0.651
P02649	*APOE; APOE3; APOE4*	0.787 – 0.894

machine. Table 3 gives details of the proteins that were collapsed into their first principal component, as detailed previously. In all cases, PC1 accounted for at least 75% of the total variation. The only probe leading to correlations of less than 0.75 for UniProt ID P01024 was that of C3d.

Discovery analysis

Discovery analysis was applied across the entire SOMAscan platform in an attempt to find predictors of dichotomized as well as continuous NAB. Variable selection techniques that included linear and logistic regression and SVMs yielded no predictive models with efficacy of estimating NAB above that of a model built on age and APOE ε4 carriage alone. The results of linear and logistic regression methods to predict continuous and dichotomized NAB are given in Supplementary Table 2.

Single protein analysis

Residual plots indicated a reasonable model fit in all cases. We outline below single proteins that pass multiple testing with a q-value of less than 0.1.

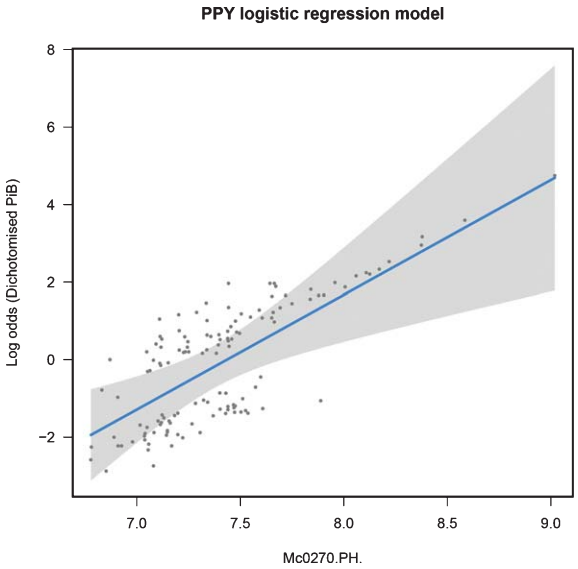


Fig. 2. Regression fit for PPY modeling dichotomized NAB in all samples: training data.

Non-AIBL candidates

When modeling dichotomized NAB, logistic regression analysis showed that PPY was significantly associated with PiB positivity across all samples, passing multiple testing corrections ($\beta = 2.959$, $p = 0.001$, $q = 0.013$). Figure. 2 shows the fit of this model in the training data. See Supplementary Table 3 for full results.

AIBL candidates

In control samples, IgM was significant in a linear model with a q-value of 0.044 ($\beta = -0.282$, $p = 0.002$). Mirroring the non-AIBL candidate results, PPY

Table 4
Multiple protein analysis: dichotomized NAB

Population	Candidates	Method	Test data					
			Accuracy	Accuracy Empirical *p*-value	Sensitivity	Sensitivity Empirical *p*-value	Specificity	Specificity Empirical *p*-value
All	Non-AIBL	Logistic regression [Age, *APOE* ε4 number, IL13, PPY, APOE PC1, C3 PC1]	0.522	0.833	0.613	0.941	0.444	0.352
		SVM	0.54	–	0.5	–	0.58	–
	AIBL	Logistic regression [Age, *APOE* ε4 number, BLC, Clusterin, PPY, C4, Hemopexin]	0.612	0.119	0.645	0.78	0.583	0.026
		SVM	0.63	–	0.59	–	0.67	–
	Age and APOE	Logistic regression [Age and *APOE* ε4 number]	0.567	–	0.774	–	0.389	–
		SVM	0.61	–	0.57	–	0.66	–
Controls	Non-AIBL	Logistic regression [Age, *APOE* ε4 number, vWF, PPY, APOE PC1, C3 PC1, Fibrinogen PC1]	0.575	0.693	0.5	0.053	0.636	0.963
		SVM	0.65	–	0.7	–	0.63	–
	AIBL	Logistic regression [Age, IgM, BLC, Alpha-2-Macroglobulin, PPY, Gelsolin, Factor B, Hemopexin]	0.675	0.083	0.444	0.217	0.864	0.155
		SVM	0.7	–	1	–	0.65	–
	Age and APOE	Logistic regression [Age and *APOE* ε4 number]	0.65	–	0.333	–	0.909	–
		SVM	0.55	–	NA**	–	0.55	–

PC1, First Principal Component;
SVM, Support Vector Machine.
**All samples are predicted to have low NAB.

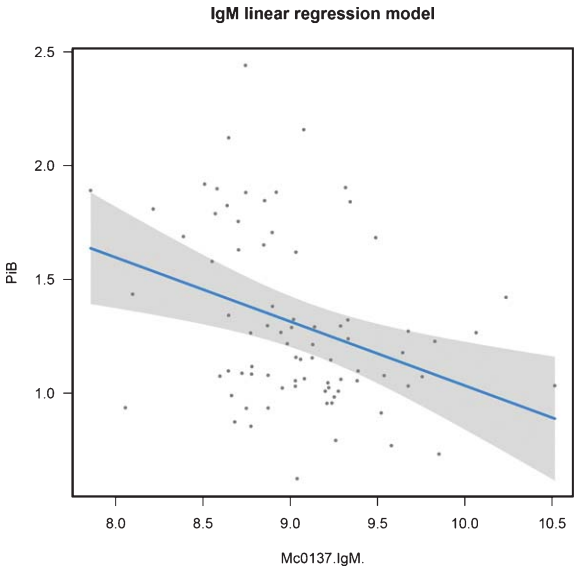


Fig. 3. Regression fit for IgM modeling continuous NAB in control samples: training data.

was significant in a logistic regression model when considering the dichotomized endpoint in all samples ($q = 0.018$). Figures 2 and 3 show the fit of these models in the training data. See Supplementary Table 4 for full results.

The analysis of continuous NAB against IgM in control subjects was repeated including a flag for presence of diabetes as a covariate ($n = 72$). IgM had an estimate similar to that previously stated ($\beta = -0.232$) and showed a nominally significant p-value ($p = 0.027$). The presence of diabetes flag was not significant ($\beta = 0.127$, $p = 0.586$).

Multiple protein analysis

Continuous NAB

In linear regression analysis across all subjects models of both non-AIBL candidates and AIBL candidates showed improved RMSE (0.490 and 0.497) and R^2 (0.245 and 0.226) values over a model of age and APOE alone (RMSE = 0.500 and $R^2 = 0.219$). The model of non-AIBL candidates achieved significant empirical p-values for RMSE and R^2 through permutation testing with p-values of 0.041 and 0.043, respectively. Model statistics achieved in parametric analysis outperformed those from SVM modeling.

Table 5
Multiple protein analysis: continuous NAB

Population	Candidates	Method	Test data			
			R^2	R^2 Empirical p-value	RMSE	RMSE Empirical p-value
All	Non-AIBL	Linear regression [Age, *APOE* ε4 number, Albumin PPY]	0.245	0.043	0.49	0.041
		SVM	0.15	–	0.52	–
	AIBL	Linear regression [Age, *APOE* ε4 number, BLC, C4]	0.226	0.061	0.497	0.058
		SVM	0.15	–	0.53	–
	Age and APOE	Linear regression [Age and *APOE* ε4 number]	0.219	–	0.5	–
		SVM	0.1	–	0.53	–
Controls	Non-AIBL	Linear regression [Age, *APOE* ε4 number, vWF, PPY, Plasmingen]	0.241	0.215	0.356	0.226
		SVM	0.23	–	0.36	–
	AIBL	Linear regression [Age, *APOE* ε4 number, IgM, BLC, Alpha-2-Macroglobulin, HRG, Factor B, Hemopexin]	0.17	0.371	0.392	0.537
		SVM	0.25	–	0.36	–
	Age and APOE	Linear regression [Age and *APOE* ε4 number]	0.273	–	0.349	–
		SVM	0.02	–	0.41	–

PC1, First Principal Component;
SVM, Support Vector Machine.

Protein models in control samples created through regression analysis did not outperform a model of age and APOE alone. However, all SVM models showed lower RMSE and higher R^2 values than the demographic only model and in most cases also outperformed the parametric methods. For further details of the multiple protein analysis of continuous NAB, see Table 5.

Dichotomized NAB

In parametric analysis across all subjects, the model of AIBL candidates gave a higher accuracy (0.612) than a model of age and APOE alone (0.567). This was driven by an increase in specificity; 0.583 compared with 0.389 in the age and APOE only model. Specificity gave a significant permutation *p*-value of 0.026.

The model of non-AIBL candidates also showed increased specificity at 0.444 compared to 0.389 in the age and APOE only model. SVM analysis gave increased specificity and decreased sensitivity over parametric analysis in all three models leading to a marginal increase in overall accuracy.

In control samples, logistic regression analysis of AIBL candidates gave improved accuracy and sensitivity compared with the model of age and APOE alone (0.675 and 0.444, respectively, compared with 0.650 and 0.333). The model in non-AIBL candidates also showed increased sensitivity at 0.500. Non-parametric analysis showed improved accuracy over regression analysis in the protein based models driven by an increase in sensitivity. For further details of the multiple protein analysis of dichotomized NAB, see Table 4.

DISCUSSION

In this study, we investigated blood proteins shown in the literature to be associated with Aβ burden in the brain, using serum samples from AIBL. No proteins were found to pass multiple testing corrections in a discovery analysis so we implemented a candidate based approach.

Two candidate proteins (PPY and IgM) showed association with NAB in the single protein analysis of the AIBL serum SOMAscan dataset. Both had already been identified as significant in the AIBL plasma RBM MAP dataset by Burnham et al. [17].

PPY passed multiple testing corrections for single protein analysis in all samples for the dichotomized endpoint. This provides further evidence for claims made by Chiam et al. who found PPY to be the most replicated blood protein marker of AD [34]. The direction of association (positive) matched that of Burnham et al. [17], providing some evidence for concordance of the SOMAscan platform with RBM MAP for this protein despite technical differences. This is also consistent with Sattlecker et al. who used SOMAscan to show that plasma PPY levels increase with disease progression in subjects from the combined AddNeuroMed, Alzheimer's Research UK/Maudsley BRC Dementia Case Registry research cohorts [18]. However, the estimate was in the opposite direction to that described by Kiddle et al. [16]. This could suggest a complex relationship of PPY with NAB that we may be unable to describe fully with the statistical approaches applied here. Alternatively, the variation in direction could be due to lower statistical power in Kiddle et al. ($n = 71$) or cohort differences, perhaps differences between Americans and Australians or disease stage. Indeed, Lunnon et al. have previously shown that disease associated markers, albeit gene expression markers, can frequently change direction of effect depending on disease stage [35]. The subjects used by Kiddle et al. from ADNI were at a much later disease stage (with only three controls) than the subjects used here.

Doecke et al. have previously shown PPY to be associated with AD diagnosis in the AIBL cohort [36]. Similarly, PPY was found to be associated with diagnosis in ADNI and the Texas Alzheimer's Research Consortium [37, 38]. However, there has been discussion around whether this effect could be a pharmacological response to the use of acetylcholinesterase inhibitors. Unfortunately, we were unable to examine this in the AIBL cohort but the possibility of medication confounding should be investigated in further studies. Chiam et al. use data from the AddNeuroMed study to suggest the effect of acetylcholinesterase inhibitor use is not significant on PPY (Chiam et al., unpublished results).

Additionally, IgM was seen to replicate at the single protein level in control samples only. IgM is an antibody produced by B-cells that appears early in the course of an infection. This pathway of immune response has been linked to AD in previous studies [35]. Furthermore, IgM has been associated with diabetes [39]. When a presence of diabetes flag is included in the IgM model it is not significant. Additionally, the *p*-value of IgM remains nominally significant at 0.027 indicating that it is unlikely that this significant association is a result of disease confounding.

There were 29 candidates that gave non-significant results. These were: C3, APOE, Fibrinogen, Hap-

toglobin, Albumin, α1-antitrypsin, α2-macroglobin, APOA1, C4, Chemokine Ligand 13, Clusterin, Factor B, Factor H, Gelsolin, Hemopexin, HRG, IgG, IL13, IL17, IL3, Leptin, MIP1α, MMP9, Plasminogen, Transferrin, SAP, VCAM1, vWF, and APP.

Some of these proteins showed significant q-values for model statistics (such as R^2 and accuracy based on permutation testing) although they were not significant in the parametric model. This indicates that greater sample size is needed for future studies of this nature in order to reduce the variation of statistics such as R^2 and hence give a more reliable estimate.

Multiple protein models constructed in this study gave mixed success in explaining variation in brain Aβ burden. Although some provided increased accuracy or R^2 over models of age and APOE ε4 number alone, the differences were minimal and largely no greater than those provided by some single protein models. LMG variable importance scores from linear regression models also emphasized that age and APOE ε4 number were the most important variables. Consequently, given a larger sample size, an analysis stratified by APOE genotype could be used to reveal more complex relationships between APOE and protein markers.

The four studies that contributed candidates for this analysis used blood plasma samples. However, due to sample availability, we performed SOMAscan analysis on blood serum samples. It is possible that this tissue difference is responsible for the non-replication of some of the candidates; however, there are studies suggesting analyte specific consistency across sample type. O'Bryant et al. show that some blood-based markers of AD, including pancreatic polypeptide, are consistent between plasma and serum [38]. Furthermore, our comparison of ranked lists of proteins associated with age between blood serum (AIBL) and plasma (Twins UK) showed a high level of concordance on the SOMAscan platform.

Differences between the proteomic approaches used here (SOMAscan) and in the literature (Myriad RBM, 2DGE, and mass spectrometry) could be driving the disparity in results between these studies. We see non-significant *p*-values for similarity when comparing ranked lists of proteins associated with age and gender between SOMAscan and RBM MAP. Due to the similarity in signal between the Twins UK plasma samples and AIBL serum samples (detailed above), this difference is likely platform-driven. While SOMAscan measures the availability of a 3D shape and charge epitope, mass spectrometry measures mass to charge ratio. RBM MAP uses antibodies that could be measuring alternative versions or conformations of proteins to those measured by SOMAscan. It is likely the SOMAscan platform would show the highest correlation with other immunoassays. This is consistent with the observation that we see the most agreement between AIBL plasma RBM MAP candidates and our findings here using serum SOMAscan (PPY and IgM). A study examining the similarities and differences observed when running samples across these platforms would be useful when interpreting results such as these.

The primary analysis used in this study follows a parametric, linear approach following that of three out of the four studies cited. Burnham et al. used a non-linear, non-parametric analysis. We investigated the potential ability to detect interactions and non-linear relationships, and consequently improve model performance, by building SVM models. These non-parametric models found slightly improved results suggesting that non-linearity and interactions could be important in the relationship between blood proteomics and NAB and should be considered in further studies.

Throughout the analysis presented here we included APOE ε4 number (0, 1, 2) as a covariate. We did not include the APOE genotype itself due to small numbers of samples in the ε2ε4 and ε4ε4 groups. We investigated whether grouping subjects with an ε2ε4 genotype with subjects with an ε3ε4 genotype was appropriate given that the ε2 allele has been found to be protective and consequently subjects of the two genotypes could behave differently. Exploratory analysis of the PPY and IgM results suggest that the general trend between the ε3ε4 and ε2ε4 groups is similar and consequently it is reasonable to combine them in our analysis. Supplementary Figures 2 and 3 illustrate this.

It is clear that the sample size used in our analysis limits the power of any conclusions. An additional factor limiting interpretation of these findings is how representative the cohort used here is compared to potential clinical trial recruitment populations, both in terms of diagnostic groups and prevalence of high Aβ burden. Differences in prevalence between this cohort and the target populations can lead to inflated estimates of positive predictive value, the proportion of test positives that are true positives [40]. This should be addressed in further work through study of larger and more representative cohorts.

A known disadvantage of stepwise regression analysis is a tendency to over fit to the training data. Several of the multiple protein models created failed to outperform age and APOE ε4 number in test data which could be a suggestion that the stepwise models are over fitting to the training data. Alternatively, this may provide evidence that the candidates considered here provide

minimal predictive information on NAB, above that of age and APOE $\varepsilon 4$ load, when measured in serum samples using SOMAscan. Studies of larger cohorts will allow this to be investigated further.

It is promising that we show for the first time that IgM predicts NAB in asymptomatic individuals. However, we note that the majority of candidates investigated in this study had been selected from populations with a variety of diagnoses. Thus, the lack of significance of association with NAB for most candidates in asymptomatic subjects was not unexpected. In further work it would be interesting to perform a discovery analysis on control subjects alone and identify proteins that are found to link with Aβ burden pre-symptomatically. The proteins identified by Thambisetty et al. are from a control population but of small sample size ($n = 57$) [15].

We were surprised to detect a signal for fibrinogen from the SOMAscan technology. Currently it is thought that there should be no fibrinogen present in blood serum samples. This may reflect issues of specificity with the fibrinogen SOMAmer or the presence of low concentrations of fibrinogen in serum samples beneath detection thresholds of other assays. We note that this has been reported before in a study of tuberculosis [41]. As SOMAscan is a relatively novel protein quantification technology, more work is needed to qualify how it compares with other platforms and to identify the binding sites of SOMAmers.

There is a clear link between the development of Aβ plaques in the brain and AD but this relationship is not exclusive. That is to say, it is not always the case that high NAB indicates AD or vice versa; there are examples of elderly subjects with high Aβ burden and no cognitive impairment [42]. While this could be explained by misdiagnosis and Aβ starting a slow AD process in motion, the alternative hypothesis would be that Aβ is not causally related to late onset AD. However, while the focus of AD clinical trials is predominately anti-Aβ, it is logical to find markers relevant to those trials. Furthermore, Aβ burden is appearing in new diagnostic criteria for AD [43]. More work is needed to thoroughly understand the role of brain Aβ burden in AD to validate the use mechanistic action of Aβ, it would be interesting to study a cohort of subjects with Aβ positivity of a model of Aβ burden as a diagnostic biomarker. To investigate the prior to the development of MCI or AD. Unfortunately, the current study did not have a sufficient number of samples ($n = 5$) meeting this criterion to perform such analysis. The need for cohorts of this type is clear in order to advance understanding of the clinical relevance of Aβ.

If the model statistics achieved in this study are shown to be robust across relevant, larger, and independent populations, these models could be beneficial in acting as a cost-effective enrichment filter between the general population and those predicted to have high NAB. Using a relevant simple blood test as a pre-screening tool could reduce the PET and CSF screening failure rate and hence the cost of trials.

CONCLUSIONS

In this study we investigated blood proteins previously found to be associated with Aβ burden in the brain using serum samples from AIBL and the SOMAscan proteomics technology. Two candidate proteins (PPY and IgM) showed association with NAB in the AIBL serum SOMAlogic dataset. Notably, IgM was found to associate with continuous NAB across control subjects, suggesting it may have utility for predicting Aβ-positive asymptomatic individuals. There are several reasons for a lack of significance for the other candidates including platform differences and the use of serum rather than plasma samples. To investigate the possibility of technical differences causing lack of further replication, further studies are required.

ACKNOWLEDGMENTS

We thank all the participants who took part in this study and the clinicians who referred participants. The AIBL study (http://www.AIBL.csiro.au) is a collaboration between CSIRO, Edith Cowan University (ECU), The Florey Institute of Neuroscience and Mental Health (FINMH), National Ageing Research Institute (NARI) and Austin Health. It also involves support from CogState Ltd., Hollywood Private Hospital, and Sir Charles Gairdner Hospital. The study received funding support from CSIRO, the Science and Industry Endowment Fund (www.SIEF.org.au), NHMRC and Dementia Collaborative Research Centres (DCRC), Alzheimer's Australia (AA), Alzheimer's Association, and the McCusker Alzheimer's Research Foundation.

This work was supported by the Alzheimer's Society, the NIHR Biomedical Research Centre for Mental Health and Biomedical Research Unit for Dementia at the South London, Maudsley NHS Foundation Trust and Kings College London. The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement number 115372, resources of which are composed of financial contribution from the European

Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution. SOMAscan and SOMAmer are trademarks of SomaLogic, Inc.

Sally Nelson and Steven Williams are employees of SomaLogic, Inc, and the proteomic assay reported in this manuscript was performed by them using reagents supplied by SomaLogic, Inc. Steven Kiddle is supported by an MRC Career Development Award in Biostatistics (MR/L011859/1). A portion of this work was funded by GE Healthcare and Janssen R&D. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, Department of Health, GEHC and Janssen.

Authors' disclosures available online (http://j-alz.com/manuscript-disclosures/15-0020r1).

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: http://dx.doi.org/10.3233/JAD-150020.

REFERENCES

[1] Blennow K, de Leon M, Zetterberg H (2006) Alzheimer's disease. *Lancet* **368**, 387-403.

[2] Querfurth H, LaFerla F (2011) (2010) Alzheimer's disease. *N Engl J Med* **362**, 329-344. Erratum in *N Engl J Med* **364**, 588.

[3] Hardy J, Higgins G (1992) Alzheimer's disease: The amyloid cascade hypothesis. *Science* **256**, 184-185.

[4] Villemagne V, Pike K, Chételat G, Ellis K, Mulligan R, Bourgeat P, Ackermann U, Jones G, Szoeke C, Salvado O, Martins R, O'Keefe G, Mathis C, Klunk W, Ames D, Masters C, Rowe C (2011) Longitudinal assessment of Aβ and cognition in aging and Alzheimer disease. *Ann Neurol* **69**, 181-192.

[5] Jack C, Knopman D, Jagust W, Petersen R, Weiner M, Aisen P, Shaw L, Vemuri P, Wiste H, Weigand S, Lesnick T, Pankratz V, Donohue M, Trojanowski J (2013) Personal view- tracking pathophysiological processes in Alzheimer's disease: An updated hypothetical model of dynamic biomarkers. *Lancet Neurol* **12**, 207-216.

[6] Corbett A, Ballard C (2012) New and emerging treatments for Alzheimer's disease. *Expert Opin Emerg Drugs* **17**, 147-156.

[7] Doody R, Raman R, Farlow M, Iwatsubo T, Vellas B, Joffe S, Kieburtz K, He F, Sun X, Thomas R, Aisen P, Siemers E, Sethuraman G, Mohs R (2013) A phase 3 trial of semagacestat for treatment of Alzheimer's disease. *N Engl J Med* **369**, 341–350.

[8] Doody R, Thomas R, Farlow M, Iwatsubo T, Vellas B, Joffe S, Kieburtz K, Raman R, Sun X, Aisen P, Siemers E, Liu-Seifert H, Mohs R (2014) Phase 3 trials of solanezumab for mild-to-moderate Alzheimer's disease. *N Engl J Med* **370**, 311-321.

[9] Salloway S, Sperling R, Fox N, Blennow K, Klunk W, Raskind M, Sabbagh M, Honig L, Porsteinsson A, Ferris S, Reichert M, Ketter N, Nejadnik B, Guenzler V, Miloslavsky M, Wang D, Lu Y, Lull J, Tudor I, Liu E, Grundman M, Yuen E, Black R, Brashear H (2014) Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. *N Engl J Med* **370**, 322-333.

[10] Shneider L, Mangialasche F, Andreasen N, Feldman H, Giacobini E, Jones R, Mantua V, Mecocci P, Pani L, Winblad B, Kivipelto M (2014) Clinical trials and late-stage drug development for Alzheimer's disease: An appraisal from 1984 to 2014. *J Intern Med* **275**, 251-283.

[11] Aisen P, Vellas B, Hampel H (2013) Moving towards early clinical trials for amyloid-targeted therapy in Alzheimer's disease. *Nat Rev Drug Discov* **12**, 324-324.

[12] Bazenet C, Lovestone S (2012) Plasma biomarkers for Alzheimer's disease: Much needed but tough to find. *Biomark Med* **6**, 441-454.

[13] Perrin R, Fagan A, Holtzman D (2009) Multimodal techniques for diagnosis and prognosis of Alzheimer's disease. *Nature* **461**, 916-922.

[14] Koyama A, Okereke O, Yang T, Blacker D, Selkoe D, Grodstein F (2012) Plasma amyloid-β as a predictor of dementia and cognitive decline. *Arch Neurol* **69**, 824-831.

[15] Thambisetty M, Tripaldi R, Riddoch-Contreas J, Hye A, An Y, Campbell J, Sojkova J, Kinsey A, Lynham S, Zhou Y, Ferrucci L, Wong D, Lovestone S, Resnick S (2010) Proteome-based plasma markers of brain amyloid-beta deposition in non-demented older individuals. *J Alzheimers Dis* **22**, 1099-1109.

[16] Kiddle S, Thambisetty M, Simmons A, Riddoch-Contreras J, Hye A, Westman E, Pike I, Ward M, Johnston C, Lupton M, Lunnon K, Soininen H, Kloszewska I, Tsolaki M, Vellas B, Mecocci P, Lovestone S, Newhouse S, Dobson R (2012) Plasma based markers of [11C] PiB-PET brain amyloid burden. *PLoS One* **7**, e44260.

[17] Burnham S, Faux N, Wilson W, Laws S, Ames D, Bedo J, Bush A, Doecke J, Ellis K, Head R, Jones G, Kiiveri H, Martins R, Rembach A, Rowe C, Salvado O, Macaulay S, Masters C, Villemagne V (2014) A blood-based predictor for neocortical Aβ burden in Alzheimer's disease: Results from the AIBL study. *Mol Psychiatry* **19**, 519-526.

[18] Sattlecker M, Kiddle S, Newhouse S, Proitsi P, Nelson S, Williams S, Johnston C, Killick R, Simmons A, Westman E, Hodges A, Soininen H, Kłoszewska I, Mecocci P, Tsolaki M, Vellas B, Lovestone S, Dobson R, Consortium t A (2014) Alzheimer's disease biomarker discovery using SOMAscan multiplexed protein technology. *Alzheimers Dement* **10**, 724-734.

[19] Ashton N, Kiddle S, Graf J, Ward M, Baird A, Hye A, Westwood S, Wong K, Dobson R, Rabinovici G, Miller B, Rosen H, Torres A, Zhang Z, Thurfjell L, Covin A, Hehir C, Baker D, Bazenet C, Lovestone S, The AIBL, Research, Group (2015) Blood protein predictors of brain amyloid for enrichment in clinical trials? *Alzheimers Dement (Amst)* **1**, 48-60.

[20] Gold L, Ayers D, Bertino J, Bock C, Bock A, Brody E, Carter J, Dalby A, Eaton B, Fitzwater T, Flather D, Forbes A, Foreman T, Fowler C, Gawande B, Goss M, Gunn M, Gupta S, Halladay D, Heil J, Heilig J, Hicke B, Husar G, Janjic N, Jarvis T, Jennings S, Katilius E, Keeney T, Kim N, Koch T, Kraemer S, Kroiss L, Le N, Levine D, Lindsey W, Lollo B, Mayfield W, Mehan M, Mehler R, Nelson S, Nelson M, Nieuwlandt D, Nikrad M, Ochsner U, Ostroff R, Otis M, Parker T, Pietrasiewicz S, Resnicow D, Rohloff J, Sanders G, Sattin S, Schneider D, Singer B, Stanton M, Sterkel A, Stewart A, Stratford S, Vaught J, Vrkljan M, Walker J, Watrobka M, Waugh S, Weiss A, Wilcox S, Wolfson A, Wolk S, Zhang C,

Zichi D (2010) Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS One* **5**, e15004.

[21] Zhao X, Lejnine S, Spond J, Zhang C, Ramaraj T, Holder D, Dai H, Weiner R, Laterza O (2015) A candidate plasma protein classifier to identify Alzheimer's disease. *J Alzheimers Dis* **43**, 549-563.

[22] O'Bryant S, Gupta V, Henriksen K, Edwards M, Jeromin A, Lista S, Bazenet C, Soares H, Lovestone S, Hampel H, Montine T, Blennow K, Foroud T, Carrillo M, Graff-Radford N, Laske C, Breteler M, Shaw L, Trojanowski J, Schupf N, Rissman R, Fagan A, Oberoi P, Umek R, Weiner M, Grammas P, Posner H, Martins R and for the STAR-B and BBBIG working groups (2014) Guidelines for the standardization of preanalytic variables for blood-based biomarker studies in Alzheimer's disease research. *Alzheimers Dement*, doi: 10.1016/j.jalz.2014.08.099.

[23] Kraemer S, Vaught J, Bock C, Gold L, Katilius E, Keeney T, Kim N, Saccomano N, Wilcox S, Zichi D, Sanders G (2011) From SOMAmer-based biomarker discovery to diagnostic and clinical applications: A SOMAmer-based, streamlined multiplex proteomic assay. *PLoS One* **6**, e26332.

[24] Rowe C, Ellis K, Rimajova M, Bourgeat P, Pike K, Jones G, Fripp J, Tochon-Danguy H, Morandeau L, O'Keefe G, Price R, Raniga P, Robins P, Acosta O, Lenzo N, Szoeke C, Salvado O, Head R, Martins R, Masters C, Ames D, Villemagne V (2010) Amyloid imaging results from the Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging. *Neurobiol Aging* **31**, 1275-1283.

[25] Lopresti B, Klunk W, Mathis C, Hoge J, Ziolko S, Lu X, Meltzer C, Schimmel K, Tsopelas N, DeKosky S, Price J (2005) Simplified quantification of Pittsburgh compound B amyloid imaging PET studies: A comparative analysis. *J Nucl Med* **46**, 1959-1972.

[26] R Core Team. (2014) R: *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

[27] Menni C, Kiddle SJ, Mangino M, Viñuela A, Psatha M, Steves C, Sattlecker M, Buil A, Newhouse S, Nelson S, Williams S, Voyle N, Soininen H, Kloszewska I, Mecocci P, Tsolaki M, Vellas B, Lovestone S, Spector TD, Dobson R, Valdes AM (2014) Circulating proteomic signatures of chronological age. *J Gerontol A Biol Sci Med Sci*, doi:10.1093/gerona/glu121.

[28] Yang X, Scheid S, Lottaz, C (2008) *OrderedList: Similarities of Ordered Gene Lists*. R package version 1.36.0.

[29] Altman D, Royston P (2006) Statistics Notes: The cost of dichotomising continuous variables. *BMJ Open* **332**, 1080-1080.

[30] Kuhn J, Contributions from Wing J, Weston S, Williams A, Keefer C, Engelhardt A, Cooper T, Mayer Z, the R Core Team, Benesty M, Lescarbeau R, Ziem A, Scrucca L (2014) *caret: Classification and Regression Training*. R package version 6.0-35.

[31] Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. *Ann Stat* **29**, 1165-1188.

[32] Grömping U (2006) Relative importance for linear regression in R: The package relaimpo. *J Stat Softw* **17**, 1-27.

[33] Karatzoglou A, Smola A, Hornik K Zeileis A (2004) kernlab – An S4 package for kernel methods in R. *J -Stat Softw* **11**, 1-20.

[34] Chiam J, Dobson R, Kiddle S, Sattlecker M (2015) Are blood-based protein biomarkers for Alzheimer's disease also involved in other brain disorders? A systematic review. *J Alzheimers Dis* **43**, 303-314.

[35] Lunnon K, Ibrahim Z, Proitsi P, Lourdusamy A (2012) Mitochondrial dysfunction and immune activation are detectable in early Alzheimer's disease blood. *J Alzheimers Dis* **30**, 685-710.

[36] Doecke J, Laws S, Faux N, Wilson W, Burnham S, Lam C, Mondal A, Bedo J, Bush A, Brown B, De Ruyck K, Ellis K, Fowler C, Gupta V, Head R, Macaulay S, Pertile K, Rowe C, Rembach A, Rodrigues M, Rumble R, Szoeke C, Taddei K, Taddei T, Trounson B, Ames D, Masters C, Martins R, the Alzheimer's Disease Neuroimaging Initiative and Australian Imaging Biomarker and Lifestyle Research, Group (2012) Blood-based protein biomarkers for diagnosis of Alzheimer disease. *Arch Neurol* **69**, 1318.

[37] Soares H, Potter W, Pickering E, Kuhn M, Immermann F, Shera D, Ferm M, Dean R, Simon A, Swenson F, Siuciak J, Kaplow J, Thambisetty M, Zagouras P, Koroshetz W, Wan H, Trojanowski J, Shaw L and the Biomarkers Consortium Alzheimer's Disease Plasma Proteomics Project (2012) Biomarkers associated with the apolipoprotein E genotype and Alzheimer disease. *Arch Neurol* **69**, 1310-1317.

[38] O'Bryant SE, Xiao G, Barber R, Huebinger R, Wilhelmsen K, Edwards M, Graff-Radford N, Doody R, Diaz-Arrastia R, for the Texas Alzheimer's Research & Care Consortium, for the Alzheimer's Disease Neuroimaging, Initiative (2011) A blood-based screening tool for Alzheimer's disease that spans serum and plasma: Findings from TARC and ADNI. *PLoS One* **6**, e28092.

[39] Decraene, T., Vandewalle, C., Pipeleers, D., Gorus, F. and the Belgian Diabetes Registry (1992) Increased concentrations of total IgM at clinical onset of type 1 (insulin-dependent) diabetes: Correlation with IgM binding to cells. *Clin Chem* **38**, 1762-767.

[40] Mayeux, R (2004) Biomarkers: Potential uses and limitations. *NeuroRx* **1**, 182-188.

[41] De Groote M, Nahid P, Jarlsberg L, Johnson J, Weiner M, Muzanyi G, Janjic N, Sterling D, Ochsner U (2013) Elucidating novel serum biomarkers associated with pulmonary tuberculosis treatment. *PLoS One* **8**, e61002.

[42] Snowdon D (1997) Aging and Alzheimer's disease: Lessons from the Nun study. *Gerontologist* **37**, 150-156.

[43] Dubois B, Feldman H, Jacova C, Hampel H, Molinuevo J, Blennow K, DeKosky S, Gauthier S, Selkoe D, Bateman R, Cappa S, Crutch S, Engelborghs S, Frisoni G, Fox N, Galasko D, Habert M, Jicha G, Nordberg A, Pasquier F, Rabinovici G, Robert P, Rowe C, Salloway S, Sarazin M, Epelbaum S, de Souza L, Vellas B, Visser P, Schneider L, Stern Y, Scheltens P, Cummings J (2014) Advancing research diagnostic criteria for Alzheimer's disease: The IWG-2 criteria. *Lancet Neurol* **13**, 614-629.

Chapter 3

A pathway based classification method for analyzing gene expression for AD diagnosis

Journal of Alzheimer's Disease 49 (2016) 659–669
DOI 10.3233/JAD-150440
IOS Press

A Pathway Based Classification Method for Analyzing Gene Expression for Alzheimer's Disease Diagnosis

Nicola Voyle[a,b,*], Aoife Keohane[a], Stephen Newhouse[a,e], Katie Lunnon[c], Caroline Johnston[a,e], Hilkka Soininen[d], Iwona Kloszewska[f], Patrizia Mecocci[g], Magda Tsolaki[h], Bruno Vellas[i], Simon Lovestone[a,j], on behalf of the AddNeuroMed consortium[k], Angela Hodges[a], Steven Kiddle[a,b,1] and Richard J.B. Dobson[a,e,1,*]

[a]*Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK*
[b]*MRC Social, Genetic and Developmental Psychiatry Centre, King's College London, London, UK*
[c]*University of Exeter Medical School, Exeter, UK*
[d]*Department of Neurology, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland*
[e]*NIHR Biomedical Research Centre for Mental Health and Biomedical Research Unit for Dementia at South London and Maudsley NHS Foundation, London, UK*
[f]*Medical University of Lodz, Lodz, Poland*
[g]*Institute of Gerontology and Geriatrics, University of Perugia, Perugia, Italy*
[h]*3rd Department of Neurology, Aristotle University, Thessaloniki, Greece*
[i]*INSERM University of Toulouse, Toulouse, France*
[j]*Department of Pyschiatry, Oxford University, Oxford, UK*
[k]*http://www.innomed-addneuromed.com*

Handling Associate Editor: Gary Arendash

Accepted 31 August 2015

Abstract.
Background: Recent studies indicate that gene expression levels in blood may be able to differentiate subjects with Alzheimer's disease (AD) from normal elderly controls and mild cognitively impaired (MCI) subjects. However, there is limited replicability at the single marker level. A pathway-based interpretation of gene expression may prove more robust.
Objectives: This study aimed to investigate whether a case/control classification model built on pathway level data was more robust than a gene level model and may consequently perform better in test data. The study used two batches of gene expression data from the AddNeuroMed (ANM) and Dementia Case Registry (DCR) cohorts.
Methods: Our study used Illumina Human HT-12 Expression BeadChips to collect gene expression from blood samples. Random forest modeling with recursive feature elimination was used to predict case/control status. Age and APOE ε4 status were used as covariates for all analysis.
Results: Gene and pathway level models performed similarly to each other and to a model based on demographic information only.

[1]Joint last authors.

*Correspondence to: Nicola Voyle and Richard Dobson, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK. Tel.: +020 7848 0924; Fax: +020 7848 0866; E-mails: nicola.voyle@kcl.ac.uk; richard.j.dobson@kcl.ac.uk

Conclusions: Any potential increase in concordance from the novel pathway level approach used here has not lead to a greater predictive ability in these datasets. However, we have only tested one method for creating pathway level scores. Further, we have been able to benchmark pathways against genes in datasets that had been extensively harmonized. Further work should focus on the use of alternative methods for creating pathway level scores, in particular those that incorporate pathway topology, and the use of an endophenotype based approach.

Keywords: Alzheimer's disease, blood, gene expression, pathways

INTRODUCTION

The most common form of dementia is Alzheimer's disease (AD). It is predicted that by 2050, 1 in every 85 people will be living with the disease [1]. No disease modifying treatments are available for AD and existing treatments only provide short-term symptomatic relief in a subset of patients [2]. Additionally, in the early stages (between 2 and 15 years prior to the development of clinical symptoms) the disease is difficult to diagnose. Villemagne et al. and Jack et al. hypothesize that characteristic AD pathology (the presence of amyloid-β (Aβ) plaques and hyperphosphorylated tau tangles in the brain) begins to develop up to 20 years prior to clinical diagnosis [3, 4]. This extended prodromal stage is an important window in which to target treatments that may be able to alter the course of the disease; provided people could be sensitively and accurately diagnosed. Aβ, tau, and phosphorylated-tau levels are indicative of AD pathology in this prodromal period and can be measured in cerebrospinal fluid (CSF) and by positron emission tomography (PET) imaging [5]. The procedures involved in attaining these measurements can be invasive or expensive and require specialized administration, equipment. and expertise. The development of a less invasive, potentially cheaper technique, such as a blood test, would offer significant advantages [6].

Recent studies indicate that gene expression levels in blood may be able to differentiate AD subjects from normal elderly controls and mild cognitive impairment (MCI) subjects with prodromal disease [7–10]. Han et al. provide an overview of studies of gene expression associated with AD-related phenotypes [11]. They state that the blood transcriptome is vital in the disease mechanism of AD and should therefore be investigated further in independent studies of a large sample size. A more general summary of gene expression data in neurodegenerative diseases is given by Cooper-Knock et al. [12]. This review emphasizes the dysregulation in neuroinflammation and intracellular signaling pathways including calcium signaling in AD. The commonality between these reviews is that they both highlight limited replicability at the single marker level. Furthermore, Han et al. report a greater concordance between differentially expressed genes at the pathway level. A pathway-based interpretation of gene expression may therefore prove more robust across different sample populations. Such an approach may also reduce noise and dimensionality. It is important to note that differential gene expression, as described in these reviews, does not necessarily identify genes that will be useful in a classification context.

Although previous gene expression studies in AD have retrospectively identified pathways altered in disease [9], this is the first study to use pathway scores for each individual to build predictive models across the population. This study used Pathway Level Analysis of Gene Expression (PLAGE) to estimate pathway variability across samples in the population by calculating sample-wise pathway scores [14]. PLAGE outperformed other single sample enrichment methods such as ZSCORE, Gene Set Variation Analysis (GSVA), and Single Sample Gene Set Enrichment Analysis (SSGSEA) in a comparison of sensitivity, specificity, and prioritization by Tarca et al. [15]. PLAGE scores have been used in univariate t-testing and unsupervised clustering methods to investigate the pathways involved in oral leukoplakia and those leading to cell proliferation and migration in leukemia [16, 17]. We combine, for the first time, PLAGE scoring with a supervised machine learning approach to build an AD classifier.

This study used blood expression data from subjects participating in the AddNeuroMed (ANM) and Dementia Case Registry (DCR) studies to develop models of clinical diagnosis. The performance metrics of gene expression and demographic models is compared with those generated using pathway level measures of expression.

MATERIALS AND METHODS

Cohort

ANM is a European multi-center study aiming to develop biomarkers for AD [18]. Subjects with an AD diagnosis as well as those with MCI and healthy con-

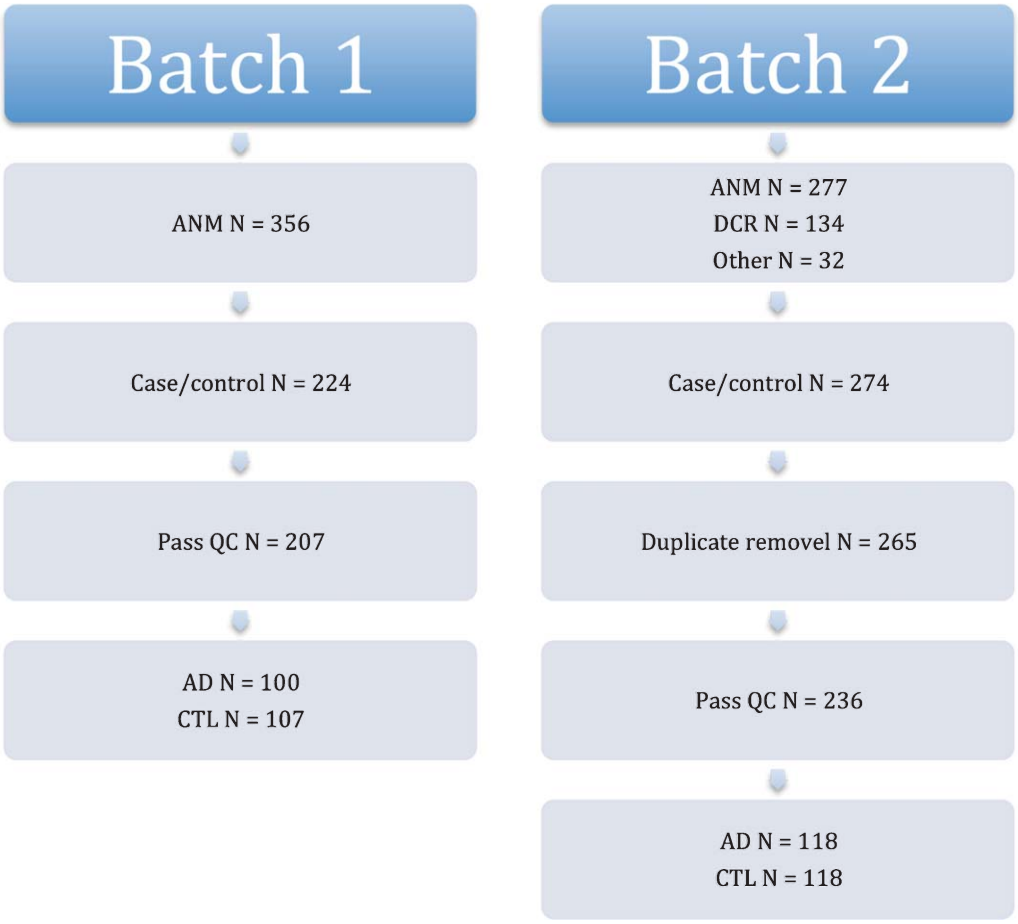


Fig. 1. Overview of sample numbers in batch 1 and 2 gene expression.

trols were recruited from centers based in Kuopio, Lodz, London, Perugia, Thessaloniki, and Toulouse. Details of study design and enrollment are provided by Lunnon et al. [10]. Subjects for the DCR were recruited from the Maudsley and Kings Healthcare Partners, which incorporates the Alzheimer's Research UK (ARUK) cohort [19] from whom gene expression data has not previously been reported.

The present study used data from 748 subjects: 614 subjects from ANM and 134 subjects from DCR.

Gene expression

Whole blood samples (2.5 ml) were collected after 2 h of fasting into Paxgene Blood RNA tubes (BD) and extracted as in Lunnon et al. [9]. Illumina Human HT-12 Expression BeadChips were used to analyze the whole transcriptome according to the manufacturers protocol. The gene expression analysis was run in two batches at two different sites. Batch 1 contained samples from 356 ANM subjects run on version 3 of the BeadChip, as previously described [9, 10]. Batch 2 contained samples from 411 subjects: 134 from DCR and 277 from ANM run on version 4 of the BeadChip. Samples from 19 subjects were included in both batches. See Fig. 1 for an overview of sample numbers. The raw gene expression data are available as GEO DataSets (Accession number GSE63060 for batch 1 and GSE63061 for batch 2).

Statistical analysis

Data pre-processing

The data pre-processing performed in this study is different to that used for the original analysis by Lunnon et al. [9, 10]. The data processing pipeline used in this study aims to address the effects of technical data artifacts in gene expression studies [20]. Raw gene expression data was exported from Illumina's Genome studio and processed in R (version 3.1.1) [21] using the

lumi package [22] and custom in-house pre-processing scripts (GitHub, http://bit.ly/1vjyKNo). Briefly, raw expression data was subject to a model based background correction for bead array [23]. This used negative bead expression levels to correct for background noise. The data was then log base 2 transformed and robust spline normalized in lumi [22]. Outlying samples were iteratively identified using fundamental network concepts and removed, following the methods described by Oldham et al. [24]. To reduce any batch effects we adjusted for technical categorical variables using ComBat [25]. Continuous technical artifacts were accounted for by taking the first principal component across housekeeping and undetected probes and regressing this against technical variables. Variables significantly associated with the first principal component were then regressed against expression for each probe, and the mean adjusted residuals taken forward for all further analyses. Finally, the data was reduced to a subset of probes that could be reliably detected in 80% of samples in at least one diagnostic group. Finally, subjects were excluded where there were discrepancies between the recorded sex and sex determined by the XIST (ILMN 1764573), USP9Y (ILMN 2056795) and EIF1AY (ILMN 1755537 and ILMN 2228976) X- and Y-linked genes.

Demographic data for the ANM and DCR subjects was extracted using CohortExplorer [26].

Pathway level analysis of gene expression (PLAGE)

Gene level expression data were condensed to sample wise, pathway level scores using PLAGE [14]. PLAGE groups genes into pathways defined by the Broad Institute Collection of Curated Pathways [27] and outputs a score, per sample, for each of these sets. We restricted PLAGE to only include pathways with between 10 and 500 genes. The generation of PLAGE scores was implemented through R package 'GSVA' and is detailed in Supplementary Methods, section 1 [13].

Data analysis

Clinical diagnosis (AD versus non-demented elderly control) classification models were built using batch 1 gene expression data. Variable selection was performed using recursive feature elimination (RFE) and the creation of a tolerance set using the 'pickSizeTolerance' function in R. This function finds a smaller set of variables while maintaining model accuracy [28]. Three Random Forest (RF) models were built, the first

Table 1
Population demographics

	AD	Control	*p*-value
Batch 1			
n	100	107	
Sex (% female)	69	58.9	0.149
APOE status (% of APOE ε4 positive)	57	32.7	<0.001
APOE ε4 load (% with loads 0; 1; 2)	43; 40; 17	67.3; 29; 3.7	<0.001
Median age [IQR] (years)	76 [10]	73 [9]	<0.001
Median MMSE score [IQR]	22 [7.25]	29 [1]	<0.001
Median years in fulltime education [IQR]	7 [5]	11 [8]	<0.001
Sample collection site			0.011
(% from KPO; LDZ; LND;	32; 15; 7;	21.5; 13.1; 21.5;	
PRG; THS; TLS)	26; 12; 8	21.5; 6.5; 15.9	
Batch 2			
n	118	118	
Sex (% female)	63.6	61.9	0.893
APOE status (% of APOE ε4 positive)	52.5	24.6	<0.001
APOE ε4 load (% with loads 0; 1; 2)	47.5; 39.8; 12.7	75.4; 20.3; 4.2	<0.001
Median age [IQR] (years)	78 [9]	74 [8]	0.001
Median MMSE score [IQR]	21 [8]	29 [2]	<0.001
Median years in fulltime education [IQR]	9 [7]	11 [5]	0.001
Sample collection site			0.002
(% from KPO; LDZ; LND;	10.2; 18.6; 35.6;	17.8; 7.6; 51.7;	
PRG; THS; TLS)	19.5; 10.2; 5.9	17.8; 3.4; 1.7	

Individuals were positive for APOE ε4 if at least one APOE ε4 allele was seen in their genotype. APOE ε4 load was the number of alleles seen in a subjects genotype. Kruskal Wallis Chi-Squared was used to test between cases and controls for continuous data. Fishers exact was used to test between cases and controls for categorical data. KPO, Kuopio; LDZ, Lodz; LND, London; PRG, Perugia; THS, Thessaloniki; TLS, Toulouse.

of which was a model based on demographic data alone (*demographic model*) [29]. The demographic variables included were those that were significant in the batch 1 population: sample collection site, age, years in full time education, and APOE status (defined as the presence of any number of $\varepsilon 4$ alleles) (Table 1). Two further models were built based on these demographic variables and gene level data (*gene model*) or PLAGE scores (*pathway model*). The purpose of the *demographic model* is to provide a comparator for the gene and pathway models. If models that include blood expression information (as well as demographics) are no more informative than demographic variables alone there is no benefit in including this information. All model building was performed in the statistical software R (Version 3.1.1) using the 'caret' package [28].

Each model was used to predict the diagnostic status of subjects in batch 2. Model statistics including accuracy, sensitivity, and specificity were generated and compared between the *demographic model*, *gene model*, and *pathway model*. Receiver Operator Curve (ROC) analysis was also performed in batch 2 data using R packages ROCR and pROC [30, 31].

Full details of model building are provided in Supplementary Methods section 2.

Additionally, variable importance (determined as the change in Gini index) was examined in the *pathway model* by permutation testing. The idea of permutation testing is to break the association between outcome (in this case diagnosis) and predictor variables. When the model is re-built based on this permuted data any significant association is spurious. Therefore, by comparing the true variable importance of a pathway to the variable importances that arise by chance in the permuted data we can assess how significant our result is. To achieve this here we used 1000 permutations of the demographic variables (including diagnosis) and for each permutation built a RF model. The importance measures of each pathway were then compared to that of the original model to generate an empirical *p*-value. A *p*-value of less than 0.05 was considered significant.

The validity of the pathways selected in the *pathway model* was also investigated in a similar way. A random set of pathways (of the same size as the final *pathway model*) were selected, and used to build a RF model. The fact that this set is random breaks the association between predictor variables and outcome. This process was repeated 1000 times and the accuracies across all models compared to the accuracy of the true model to create an empirical *p*-value.

RESULTS

Cohort demographics

Table 1 gives an overview of the demographics of subjects included in the two batches of gene expression data.

Data pre-processing

As a result of pre-processing 12 samples in batch 1 and 49 samples in batch 2 failed quality control (QC) and were removed. The majority of these samples failed QC as they were identified as outliers. Additionally, some samples were removed because the sex of the individual recorded in the clinical database did not match the biological sample (2 samples in batch 1 and 7 in batch 2).

Samples from 19 subjects were present in both batch 1 and batch 2. Samples from 14 of these individuals passed QC in both batches; only data from batch 1 was used and the other was discarded. Correlation between the two batches was at least 0.9 for all individuals (Supplementary Figure 3). Batch 2 gene expression data contains subjects from the DCR whereas batch 1 does not. This study used the same protocols, staff, and facilities as the London sample collection site within ANM. Principal components analysis (PCA) was performed across the batch 2 gene expression data from DCR and ANM subjects from London. The first three principal components (accounting for >40% of variation) were linearly regressed against the study the individual was enrolled in (DCR or London ANM) and found to be non-significant. Therefore, it was deemed appropriate to group DCR subjects with London ANM, allowing the model trained in batch 1 data to be simply applied to batch 2 data.

After data processing only subjects with either an AD diagnosis at all visits or control status at all visits were analyzed further: 207 subjects in batch 1 and 236 in batch 2.

Only gene probes that mapped between the version 3 and version 4 chips used to generate batches 1 and 2, respectively, were used for analysis (5212 probes). The Broad Institute Collection of Curated Pathways matched these probes to 834 pathways [27].

Data analysis

Demographic model

The following demographic variables were associated with case/control status in our cohorts (Table 1):

Table 2
Random Forest model results in independent test data

Model	Accuracy [95% CI]	Sensitivity	Specificity	AUC ROC
Demographic model	0.686 [0.623; 0.745]	0.534	0.839	0.771
Demographic model (no samples collection site)	0.674 [0.610; 0.733]	0.678	0.669	0.761
Pathway model	0.657 [0.592; 0.717]	0.610	0.703	0.729
Gene model	0.657 [0.592; 0.717]	0.568	0.746	0.724

CI, Confidence interval; AUC ROC, Areas under the receiver operating curve.

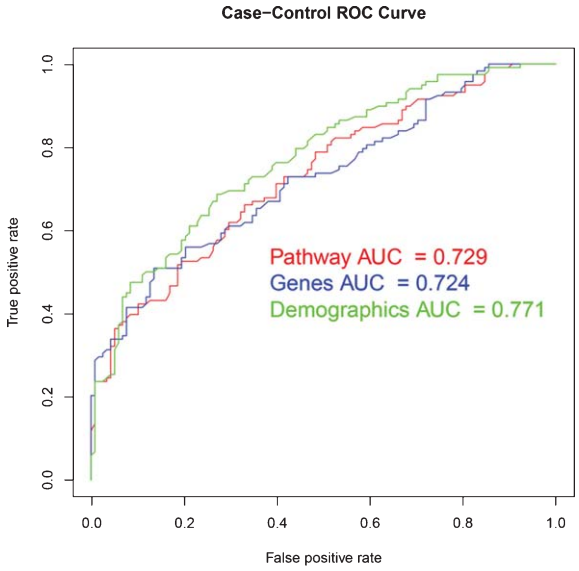


Fig. 2. ROC curves for Random Forest models in independent test data. ROC, Receiver Operating Characteristic; AUC, Area under the curve.

age, sex, APOE status, years in full time education, and sample collection site. These variables were therefore used in multivariate modeling using RFE. The optimal cross-validated accuracy was found when including all variables; calculation of a tolerance set excluded the variable representing the Lodz sample collection site. Variable importance scores showed age as the most important covariate followed by years in full time education and then APOE status and sample collection site. In batch 2 test data the model achieved an accuracy of 0.69, sensitivity of 0.53 and specificity of 0.84.

The area under the ROC curve was 0.77 (see Table 2 and Fig. 2).

Additionally, a model that did not contain the sample collection site was built. The aim was to create a model based on demographics that would be available to clinicians. This model had a slightly decreased accuracy in comparison to the *demographic model* but outperformed the *pathway model* and *gene model* in accuracy, sensitivity and area under the ROC curve at 0.67, 0.68 and 0.76, respectively. Interestingly, the specificity of the model was lower than all others at 0.67 (Table 2).

Gene model

The top 5% of variables from the bootstrapped variable importance calculations (261 variables) were carried forward to the RFE model building. The optimal cross-validated accuracy from RFE in the *gene model* was found for all of the 261 variables; calculation of a tolerance set reduced this set to only

Table 3
Genes in *gene model* with variable importance scores

Gene (Illumina ID)	Variable importance	Gene symbol	Entrez ID	Gene name
ILMN_2189936	11.9	RPL36AL	6166	Ribosomal protein L36a-like
ILMN_2189933	10.8	RPL36AL	6166	Ribosomal protein L36a-like
ILMN_2097421	10.5	MRPL51	51258	Mitochondrial ribosomal protein L51
ILMN_2237746	10.4	ING3	54556	Inhibitor of growth family, member 3
ILMN_1695645	9.2	CETN2	1069	Centrin, EF-hand protein, 2
ILMN_1784286	7.9	NDUFA1	4694	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex
ILMN_1652073	7.0	LOC653658	653658	Ribosomal protein S23 pseudogene 8
ILMN_1716053	7.0	AK2	204	Adenylate kinase 2
ILMN_1732328	6.5	LOC646200	646200	
ILMN_1776104	5.9	NDUFS5	4725	NADH dehydrogenase (ubiquinone) Fe-S protein 5
ILMN_1753892	5.8	LOC654121	654121	
ILMN_1745343	5.4	ZMAT2	153527	Zinc finger, matrin-type 2
ILMN_2048326	4.7	RPS27A	6233	Ribosomal protein S27a

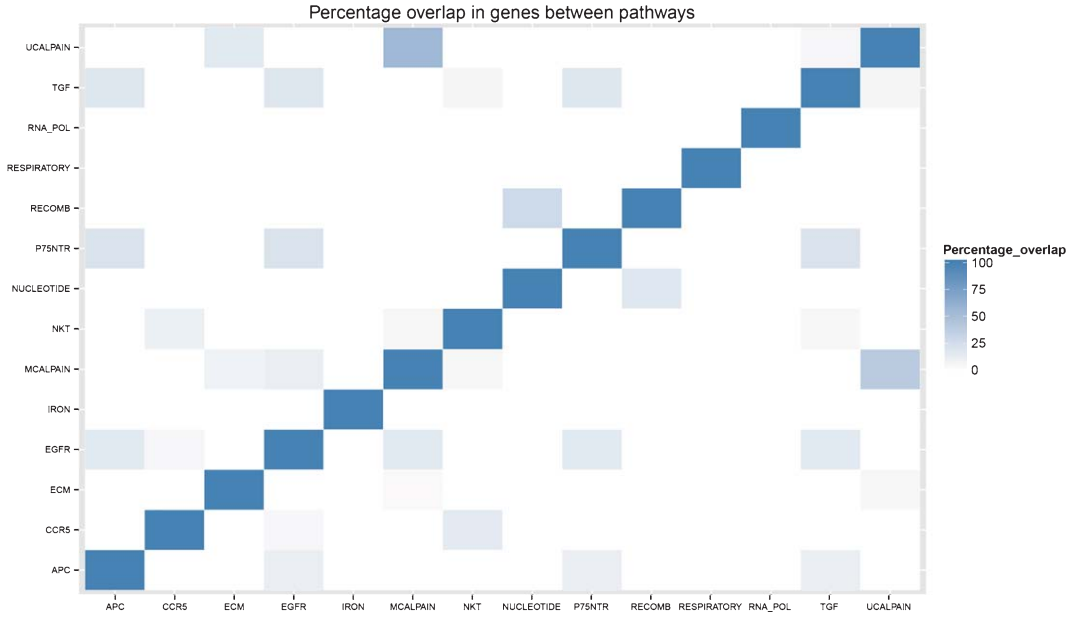


Fig. 3. Percentage overlap of genes belonging to pathways selected for the Random Forest *pathway model*.

Table 4
Pathways in *pathway model* with variable importance scores

Pathway	Abbreviation	Number of genes in pathway	Variable importance
KEGG HOMOLOGOUS RECOMBINATION	RECOMB	28	13.6*
BIOCARTA MCALPAIN PATHWAY	MCALPAIN	25	7.8*
REACTOME APC C CDC20 MEDIATED DEGRADATION OF CYCLIN B	APC	26	7.5
REACTOME TGF BETA RECEPTOR SIGNALING IN EMT EPITHELIAL TO MESENCHYMAL TRANSITION	TGF	16	7.2
REACTOME P75NTR SIGNALS VIA NFKB	P75NTR	14	6.9
BIOCARTA UCALPAIN PATHWAY	UCALPAIN	18	6.8
REACTOME RNA POL III TRANSCRIPTION	RNA_POL	33	6.4
BIOCARTA NKT PATHWAY	NKT	29	6.0
KEGG NUCLEOTIDE EXCISION REPAIR	NUCLEOTIDE	44	5.9
REACTOME IRON UPTAKE AND TRANSPORT	IRON	36	5.3
BIOCARTA CCR5 PATHWAY	CCR5	20	5.3
KEGG ECM RECEPTOR INTERACTION	ECM	84	5.0
REACTOME SIGNALING BY CONSTITUTIVELY ACTIVE EGFR	EGFR	18	5.0
REACTOME RESPIRATORY ELECTRON TRANSPORT ATP SYNTHESIS BY CHEMIOSMOTIC COUPLING AND HEAT PRODUCTION BY UNCOUPLING PROTEINS	RESPIRATORY	98	4.4

*Nominally significant in permutation testing ($p < 0.05$).

13, excluding all demographic variables. For a list of genes, see Table 3.

In batch 2 test data, the *gene model* accuracy was lower than that of the *demographic model* and equal to the *pathway model*. The sensitivity, specificity, and area under the ROC curve of the *gene model* lay between the demographic and pathway models at 0.59, 0.75, and 0.72, respectively. (see Table 2 and Fig. 2). Note that the *pathway model* showed higher sensitivity while specificity and AUC ROC were higher in the *demographic model*.

Pathway model

The top 5% of variables from the bootstrapped variable importance calculations (42 variables) were carried forward to the RFE model building. The optimal cross-validated accuracy from RFE in the *pathway model* was found for 40 of the variables; calculation of

a tolerance set reduced this set to only 16 variables (Table 4): 14 pathways, age, and years in full time education.

Permutation tests of variable importance were performed to assess the size of effect relative to that observed under the null hypothesis of no association. Of the 14 pathways, two achieved nominal significance with a p-value <0.05 and are indicated with a * in Table 4 (Supplementary Figure 5).

Additionally, we compared the model accuracy of 1000 models comprising 16 random variables from the pathways, age, sex, APOE status, and years in full time education. This yielded a p-value of 0.082 indicating that, statistically, the final model does not perform significantly better than a model of random pathways (Supplementary Figure 4).

In batch 2 test data, the model accuracy was lower than that of the *demographic model* at 0.66, however, the sensitivity was higher at 0.61. Both specificity and area under the ROC curve were lower than the *demographic model* at 0.70 and 0.73, respectively (see Table 2 and Fig. 2).

There is minimal overlap in genes between the different pathways included in the final *pathway model*. This is illustrated by the sparse percentage overlap map shown in Fig. 3 and supports the idea that each pathway is contributing an independent signal to the model. Of the 13 genes included in the *gene model*, only four of them (ILMN 1776104, ILMN 1784286, ILMN 1695645, and ILMN 2048326) appear in any of the pathways in the *pathway model*.

Misclassification

We discovered that 22% of controls used in the training data had reported memory complaints deemed not serious enough to reflect a change in diagnosis. By studying misclassification rates split by AD subjects, control subjects, and control subjects with memory complaints, we see that the most well classified group in the *gene model* was those subjects with memory complaints whereas in the *pathway model* it was control subjects (see Supplementary Figure 6). We also demonstrated that time since disease onset is not related to misclassification of AD subjects and control subjects with memory complaints in the test data (Supplementary Figure 7).

DISCUSSION

In this study we investigated whether AD cases could be differentiated from control subjects using gene expression data analyzed at the pathway level. We were particularly interested in confirming whether pathway level information created a more robust predictor of case/control status than expression data at the gene level as recent reviews of AD studies have suggested [11]. Our results, using subjects from the ANM and DCR cohorts, show similar model performance in a *pathway model* compared to a gene and demographic only model. In this study, we do not find improved prediction of AD diagnosis using pathway level information using the PLAGE method to calculate pathway scores. However, the robustness of pathway based approaches for AD biomarker discovery should be tested in other gene expression data from different populations and platforms.

The fourteen pathways included in the final *pathway model* focused around DNA repair, immune response, and regulation of cellular activities. Of particular interest to AD, two pathways containing genes from the calpain gene family were included in the final fourteen pathways. It is thought that amyloid peptides interfere with calpain activity leading to deregulation of the CDK5 gene and in turn hyperphosphorylation of the tau protein. This promotes the death of neurons [27]. It is encouraging that we have seen relevant pathways in our final *pathway model*. Overall, the pathways are similar to those identified by Lunnon et al. who studied overall pathway differences using an identical raw dataset that was processed differently [9]. As we would expect, 12 out of 13 of the genes in the final *gene model* were present in the genes used for modeling by Lunnon et al. The data had been processed slightly differently emphasizing that these signals are robust to alterations in processing and modeling methods.

RF models are commonly used in biomarker studies [9, 32, 33]. However, it has been shown that they exhibit variable selection bias being more likely to select continuous variables or those with many categories [34]. Additionally, the presence of correlated predictors (as is common in gene expression studies) can add further bias [35]. Strobl et al. aimed to address these issues with an ensemble-learning algorithm based on conditional inference trees; Conditional RF (CRF) models [36, 37]. We attempted to use this methodology in the present study. We hypothesized that the creation of an unbiased predictor may highlight different pathways and genes to those previously discovered, potentially allowing greater predictive ability. However, the process of creating a CRF model was computationally expensive even when using high performance com-

puting resources. Model building considering the 834 pathways and 5,212 genes was consequently infeasible. Work to improve the efficiency of this method would be computationally beneficial and would allow the use of alternative variable importance measures. Measures such as mean decrease in accuracy and conditional mean decrease in accuracy would be an improvement over biased variable importance measures such as the Gini index, which was used in this study.

This study used the Broad institute collection of curated pathways to generate the *pathway model* and excluded less well-curated gene sets. This method was chosen due to its performance in a comparison study [15] and due to ease of application through the GSVA R package. It may be beneficial, although potentially computationally costly, to create pathway level scores that also reflect pathway topology and thus add further detail to the model. Such methods have been created by Pyatnitskiy et al. [38] building on the work of others [15, 39]. The method detailed by Pyatnitskiy et al. does not depend on predefined gene sets as used in this analysis. However, it is also unable to control the number of genes in a pathway; a potential benefit of using PLAGE. A further limitation of existing pathway approaches is that they often ignore information on the direction of change for each gene within a pathway. This would be an interesting area for further method development.

The creation of a *demographic model* that excluded sample collection site led to a drop in accuracy. Although RNA extraction and analysis were performed at one site the blood collection may vary by location. We aimed to correct for batch effects occurring in extraction and analysis in the pre-processing. This highlights that although sample collection sites within multi-center studies are following the same protocols major technical differences can still arise and remain after QC steps including batch correction. As much as possible, these differences should be quantified during extraction. Standardization for future biomarker development will aid this. It is possible that the sample collection site effect we see is driven by genetic differences between sites for some genes (expression quantitative trait loci). For a biomarker to have clinical utility it should be robust to such differences. However, in early exploratory work we are more likely to find results of interest if technical data artifacts are not creating a barrier.

The models created in this study all achieved an accuracy of approximately 70% with the *pathway model* having test sensitivity and specificity results of greater than 60%. The *pathway model* and *gene model* did not outperform a model of demographics alone. Any potential increase in concordance from the novel pathway level approach used here has not lead to a greater predictive ability in these datasets. However, we have only tested one method for creating pathway level scores. Further, we have been able to benchmark pathways against genes in datasets that had been extensively harmonized. It is reassuring to see that pathways perform similarly to genes and further work is now needed to see if pathway concordance is more easily detected using other methodological approaches and in data generated by independent groups and platforms.

Furthermore, we found that the heterogeneity of control subjects may be leading to reduced predictive accuracy and suggest that the use of an endophenotype may be beneficial in future work.

CONCLUSIONS

We have used subjects from the ANM and DCR studies to investigate case/control classification using gene and pathway level expression data. We hypothesized that a model built on pathway level data may be more robust than a gene level model and consequently perform better in test data. However, a pathway level model built using scores and a gene level model performed similarly to each other and to a model based on demographic information only. Further work should focus on the use of alternative methods for creating pathway level scores, in particular those that incorporate pathway topology, and the use of an endophenotype based approach.

ACKNOWLEDGMENTS

This work was supported by the Alzheimer's Society, InnoMed (Innovative Medicines in Europe), an integrated project funded by the European Union of the Sixth Framework program priority (FP6-2004-LIFESCIHEALTH-5); Alzheimer's Research Trust UK; the John and Lucille van Geest Foundation (AH); and the NIHR Biomedical Research Centre for Mental Health and Biomedical Research Unit for Dementia at the South London, Maudsley NHS Foundation Trust and Kings College London, and a joint infrastructure grant from Guy's and St Thomas' Charity and the Maudsley Charity. The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under

EMIF grant agreement number 115372, resources of which are composed of financial contribution from the European Union's Seventh Framework Program (FP7/2007-2013) and EFPIA companies' in kind contribution. Kuopio University Hospital (HS) and funding from UEF- BRAIN (HS). Steven Kiddle is supported by an MRC Career Development Award in Biostatistics (MR/L011859/1).

Authors' disclosures available online (http://j-alz.com/manuscript-disclosures/15-0440r2).

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: http://dx.doi.org/10.3233/JAD-150440.

REFERENCES

[1] Brookmeyer R, Johnson E, Ziegler-Graham K, Arrighi HM (2007) Forecasting the global burden of Alzheimer's disease. *Alzheimers Dement* **3**, 186-191.

[2] Corbett A, Ballard C (2012) New and emerging treatments for Alzheimer's disease. *Expert Opin Emerg Drugs* **17**, 147-156.

[3] Villemagne V, Pike K, Chételat G, Ellis K, Mulligan R, Bourgeat P, Ackermann U, Jones G, Szoeke C, Salvado O, Martins R, O'Keefe G, Mathis C, Klunk W, Ames D, Masters C, Rowe C (2011) Longitudinal assessment of Aβ and cognition in aging and Alzheimer disease. *Ann Neurol* **69**, 181-192.

[4] Jack C, Knopman D, Jagust W, Petersen R, Weiner M, Aisen P, Shaw L, Vemuri P, Wiste H, Weigand S, Lesnick T, Pankratz V, Donohue M, Trojanowski J (2013) Tracking pathophysiological processes in Alzheimer's disease: An updated hypothetical model of dynamic biomarkers. *Lancet Neurol* **12**, 207-216.

[5] Cedazo-Minguez A, Winblad B (2010) Biomarkers for Alzheimer's disease and other forms of dementia: Clinical needs, limitations and future aspects. *Exp Gerontol* **45**, 5-14.

[6] Bazenet C, Lovestone S (2012) Plasma biomarkers for Alzheimer's disease: Much needed but tough to find. *Biomark Med* **6**, 441-454.

[7] Booij BB, Lindahl T, Wetterberg P, Skaane N, Sæbø S, Feten G, Rye PD, Kristiansen LI, Hagen N, Jensen M, Bårdsen K, Winblad B, Sharma P, Lönneborg A (2011) A gene expression pattern in blood for the early detection of Alzheimer's disease. *J Alzheimers Dis* **23**, 101-119.

[8] Rye P, Booij B, Grave G, Lindahl T, Kristiansen L, Anderson H, Horndalsveen P, Nygaard H, Naik M, Hoprekstad D, Wetterberg P, Nilsson C, Aarsland D, Sharma P, Lönneborg A (2011) A novel blood test for the early detection of Alzheimer's disease. *J Alzheimers Dis* **23**, 121-129.

[9] Lunnon K, Ibrahim Z, Proitsi P, Lourdusamy A (2012) Mitochondrial dysfunction and immune activation are detectable in early Alzheimer's disease blood. *J Alzheimers Dis* **30**, 685-710.

[10] Lunnon K, Sattlecker M, Furney S, Coppola G, Simmons A, Proitsi P, Lupton M, Lourdusamy A, Johnston C, Soininen H, Kloszewska I, Mecocci P, Tsolaki M, Vellas B, Geschwind D, Lovestone S, Dobson R, Hodges A, and the AddNeuroMed Consortium (2013) A blood gene expression marker of early Alzheimer's disease. *J Alzheimers Dis* **33**, 737-753.

[11] Han G, Wang J, Zeng F, Feng X, Yu J, Cao HY, Yi X, Zhou H, Jin LW, Duan Y, Wang YJ, Lei H (2013) Characteristic transformation of blood transcriptome in Alzheimer's disease. *J Alzheimers Dis* **35**, 373-386.

[12] Cooper-Knock J, Kirby J, Ferraiuolo L, Heath PR, Rattray M, Shaw PJ (2012) Gene expression profiling in human neurodegenerative disease. *Nat Rev Neurol* **8**, 518-530.

[13] Hänzelmann S, Castelo R, Guinney J (2013) GSVA: Gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics* **14**, 1-15.

[14] Tomfohr J, Lu J, Kepler T (2005) Pathway level analysis of gene expression using singular value decomposition. *BMC Bioinformatics* **6**, 225.

[15] Tarca A, Draghici S, Khatri P, Hassan S, Mittal P, Kim J, Kim C, Kusanovic J, Romero R (2009) A novel signaling pathway impact analysis. *Bioinformatics* **25**, 75-82.

[16] Gümüs Z, Du B, Kacker A, Boyle J, Bocker J, Mukherjee P, Subbaramaiah K, Dannenberg A, Weinstein H (2008) Effects of tobacco smoke on gene expression and cellular pathways in a cellular model of oral leukoplakia. *Cancer Prev Res (Phila)* **1**, 100-111.

[17] Deaglio S, Aydin S, Grand M, Vaisitti T, Bergui L, D'Arena G, Chiorino G, Malavasi F (2010) CD38/CD31 interactions activate genetic pathways leading to proliferation and migration in chronic lymphocytic leukemia cells. *Mol Med* **18**, 87-91.

[18] Lovestone S, Francis P, Kloszewska I, Mecocci P, Simmons A, Soininen H, Spenger C, Tsolaki M, Vellas B, Wahlund LO, Ward M, on behalf of the AddNeuroMed Consortium (2009) AddNeuroMed-The European collaboration for the discovery of novel biomarkers for Alzheimer's disease. *Ann N Y Sci* **1180**, 36-46.

[19] Hye A, Lynham S, Thambisetty M, Causevic M, Campbell J, Byers H, Hooper C, Rijsdijk F, Tabrizi S, Banner S, Shaw C, Foy C, Poppe M, Archer N, Hamilton G, Powell J, Brown R, Sham P, Ward M, Lovestone S (2006) Proteome-based plasma biomarkers for Alzheimer's disease. *Brain* **129**, 3042-3050.

[20] Parker H, Leek J (2012) The practical effect of batch on genomic prediction. *Stat Appl Genet Mol Biol* **11**, Article 10.

[21] R Core Team (2014) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

[22] Du P, Kibbe W, Lin S (2008) lumi: A pipeline for processing Illumina microarray. *Bioinformatics* **24**, 1547-1548.

[23] Xie Y (2010) *MBCB: MBCB (Model-based Background Correction for Beadarray)*. R package version 1.18.0.

[24] Oldham M, Langfelder P, Horvath S (2012) Network methods for describing sample relationships in genomic datasets: Application to Huntingdon's disease. *BMC Syst Biol* **6**, 63.

[25] Johnson WE, Li C, Rabinovic A (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118-127.

[26] Dixit A, Dobson R (2014) CohortExplorer: A generic application programming interface for entity attribute value database schemas. *JMIR Med Inform* **2**, e32.

[27] Subramanian A, Tamayo P, Mootha V, Mukherjee S, Ebert B, Gillette M, Paulovich A, Pomeroy S, Golub T, Lander E, Mesirov J (2005) Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550.

[28] Kuhn JM. Contributions from Wing J, Weston S, Williams A, Keefer C, Engelhardt A, Cooper T, Mayer Z, and the R Core

Team (2014) *Caret: Classification and Regression Training*. R package version 6.0-35.

[29] Liaw A, Wiener M (2002) Classification and Regression by randomForest. *R News* **2**, 18-22.

[30] Sing T, Sander O, Beerenwinkel N, Lengauer T (2005) ROCR: Visualizing classifier performance in R. *Bioinformatics* **21**, 3940-3941.

[31] Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Müller M (2011) pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* **12**, 77.

[32] Burnham S, Faux N, Wilson W, Laws S, Ames D, Bedo J, Bush A, Doecke J, Ellis K, Head R, Jones G, Kiiveri H, Martins R, Rembach A, Rowe C, Salvado O, Macaulay S, Masters C, Villemagne V (2014) A blood-based predictor for neocortical Aβ burden in Alzheimer's disease: Results from the AIBL study. *Mol Psychiatry* **19**, 519-526.

[33] Sattlecker M, Kiddle S, Newhouse S, Proitsi P, Nelson S, Williams S, Johnston C, Killick R, Simmons A, Westman E, Hodges A, Soininen H, Kloszewska I, Mecocci P, Tsolaki M, Vellas B, Lovestone S, Dobson R, and the AddNeuroMed Consortium (2014) Alzheimer's disease biomarker discovery using SOMAscan multiplexed protein technology. *Alzheimers Dement* **10**, 724-734.

[34] Strobl C, Boulesteix A, Zeileis A, Hothorn T (2007) Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics* **8**, 25.

[35] Meng Y, Yu Y, Cupples L, Farrer L, Lunetta K (2009) Performance of random forest when SNPs are in linkage disequilibrium. *BMC Bioinformatics* **10**, 78.

[36] Hothorn T, Hornik K, Zeileis A (2006) Unbiased recursive partitioning: A conditional inference framework. *J Comput Graph Stat* **15**, 651-674.

[37] Strobl C, Malley J, Tutz G (2009) An introduction to recursive partitioning: Rationale, application and characteristics of classification and regression trees, bagging and random forests. *Psychol Methods* **14**, 323-348.

[38] Pyatnitskiy M, Mazo I, Shkrob M, Schwartz E, Kotelnikova E (2014) Clustering Gene Expression Regulators: New Approach to Disease Subtyping. *PLoS One* **9**, e84955.

[39] Hung J, Whitfield T, Yang T, Hu Z, Weng Z, DeLisi C (2010) Identification of functional modules that correlate with phenotypic difference: The influence of network topology. *Genome Biol* **11**, R23.

Chapter 4

Blood metabolite markers of neocortical amyloid-β burden: discovery and enrichment using candidate proteins

OPEN

Citation: Transl Psychiatry (2016) **6**, e719; doi:10.1038/tp.2015.205

www.nature.com/tp

ORIGINAL ARTICLE

Blood metabolite markers of neocortical amyloid-β burden: discovery and enrichment using candidate proteins

N Voyle[1], M Kim[2], P Proitsi[3], NJ Ashton[3,4], AL Baird[3,5], C Bazenet[3,4], A Hye[3,4], S Westwood[3,5], R Chung[6], M Ward[6], GD Rabinovici[7], S Lovestone[3,5], G Breen[4], C Legido-Quigley[2], RJB Dobson[4,8], SJ Kiddle[1,8] for the Alzheimer's Disease Neuroimaging Initiative[9]

We believe this is the first study to investigate associations between blood metabolites and neocortical amyloid burden (NAB) in the search for a blood-based biomarker for Alzheimer's disease (AD). Further, we present the first multi-modal analysis of blood markers in this field. We used blood plasma samples from 91 subjects enrolled in the University of California, San Francisco Alzheimer's Disease Research Centre. Non-targeted metabolomic analysis was used to look for associations with NAB using both single and multiple metabolic feature models. Five metabolic features identified subjects with high NAB, with 72% accuracy. We were able to putatively identify four metabolites from this panel and improve the model further by adding fibrinogen gamma chain protein measures (accuracy = 79%). One of the five metabolic features was studied in the Alzheimer's Disease Neuroimaging Initiative cohort, but results were inconclusive. If replicated in larger, independent studies, these metabolic features and proteins could form the basis of a blood test with potential for enrichment of amyloid pathology in anti-amyloid trials.

Translational Psychiatry (2016) **6,** e719; doi:10.1038/tp.2015.205; published online 26 January 2016

INTRODUCTION

The most common form of dementia is Alzheimer's disease (AD), a neurodegenerative condition that leads to severe cognitive impairment in later life. Currently, the disease mechanism of AD is not comprehensively understood, and consequently no disease-modifying treatments are available. Unfortunately, current symptomatic treatments only have a moderate effect.[1] There is therefore a desperate need for a disease-modifying treatment for AD.

A definitive AD diagnosis can only be made post-mortem; however, neuropathological biomarkers (amyloid-β (Aβ) plaques and phosphorylated tau tangles) can be used to help differentiate AD from other dementias during a person's lifetime. These can be used in a clinical trial setting to ensure that all recruited participants have evidence of the target pathology. In a trial of Bapineuzumab, an anti-amyloid therapeutic, 14% of subjects had low amyloid. It was therefore unlikely that these subjects would see any benefit from the treatment. Furthermore, their involvement in that study would have reduced the statistical power of finding a treatment effect.[2] Many trials now test for elevated neocortical amyloid burden (NAB) as an eligibility requirement.

Elevated NAB is also becoming an eligibility criterion for some prevention trials, such as the A4 trial.[3] This trial aims to assess whether anti-amyloid therapeutics can delay early cognitive decline in asymptomatic individuals, a concept that has developed as a result of research showing that AD has a long prodromal stage.[4,5] The characteristic disease pathology of AD can begin to develop up to 20 years before any clinical symptoms.[6,7] This provides a window of time for a potential treatment to stop, or at least slow down, future progression of the disease.

Neuropathological biomarkers are measured by quantifying the concentrations of Aβ, tau and phosphorylated tau in the cerebrospinal fluid (CSF) or via positron emission tomography (PET) imaging. In addition, metabolites in CSF have been studied as possible biomarkers for AD and related phenotypes.[8] However, the methods used to capture this information are invasive, require specialized equipment and are often expensive and hence impractical on a large scale.

Consequently, there is a high demand for a blood-based biomarker of AD that would be easier and potentially cheaper to attain.[9] Metabolites are typically smaller than other biological molecules, and therefore have a greater chance of passing through a possibly weakened blood–brain barrier.[10] This increases the chance that blood metabolites could serve as a biomarker of AD. A review of AD biofluid metabolite studies has highlighted sphingolipid and glutamate metabolism as being altered in AD, besides the metabolism of molecules with antioxidant properties.[11] In addition, Proitsi *et al.* used a case–control study design to discover a set of long-chain cholesteryl esters associated with AD, whereas other studies have aimed to predict conversion from mild cognitive impairment (MCI) to AD.[12–14] Mapstone *et al.*[13] discovered a lipid panel from peripheral blood that

[1]MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; [2]Institute of Pharmaceutical Science, Kings College London, London, UK; [3]Department of Old Age Psychiatry, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; [4]NIHR Biomedical Research Centre for Mental Health and Biomedical Research Unit for Dementia at South London and Maudsley NHS Foundation, London, UK; [5]Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford, UK; [6]Proteomics Facility, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK and [7]Memory and Aging Center, University of California, San Francisco, San Francisco, CA, USA. Correspondence: Dr RJB Dobson or Dr SJ Kiddle, MRC Social, Genetic and Developmental Psychiatry Centre, King's College London, IoPPN, 16 De Crespigny Park, Denmark Hill, London SE5 8AF, UK.
E-mail: richard.j.dobson@kcl.ac.uk or steven.kiddle@kcl.ac.uk
[8]Joint senior authors.
[9]Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data, but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.
Received 18 August 2015; revised 19 October 2015; accepted 30 October 2015

predicted conversion from control status to amnestic MCI or AD with 90% accuracy. The panel highlighted metabolites involved in cell membrane integrity and lipids involved in cell signaling, as also suggested by Whiley *et al*.[15]

These studies should now be extended to identify markers of amyloid pathology. Such markers could then be used to enrich clinical trials with elevated NAB as an eligibility criterion. Using a blood test as a filter before a confirmatory lumbar puncture or PET scan could improve the efficiency of clinical trials by reducing the cost of recruitment.[16]

Analogous approaches have already been applied to identify genetic and protein biomarkers of NAB. A polygenic risk score trained on AD diagnosis has been shown to associate with CSF Aβ levels in the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort.[17] Similarly, multiple studies have identified potential blood protein biomarkers of NAB, as reviewed in Voyle *et al*.[18] Of particular interest are replicated markers of NAB including pancreatic polypeptide (PPY) and fibrinogen gamma chain (FGG).[16,18]

This is the first study to investigate associations between blood metabolites and NAB. Further, we present the first multi-modal analysis of blood markers in AD biomarker discovery. We consider whether a blood metabolite signal complements that of previously discovered blood protein biomarkers of NAB.

MATERIALS AND METHODS

Cohorts

UCSF. Subjects were recruited from those enrolled in the University of California, San Francisco (UCSF) AD Research Centre. Study information has been given elsewhere.[19,20] The study was approved by the UCSF and Lawrence Berkeley National Laboratory committees for human research. All subjects provided written informed consent before participating.

ADNI. ADNI is a longitudinal cohort study aiming to validate the use of biomarkers in AD clinical trials and diagnosis. Data used in the preparation of this article were obtained from the ADNI database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public–private partnership, led by Principal Investigator Michael W Weiner. The primary goal of ADNI has been to test whether biological markers and clinical and neuropsychological assessment can be combined to measure the progression of MCI and AD. For information, see www.adni-info.org. ADNI was approved by the institutional review boards of all participating institutions, and written informed consent was obtained from all participants.

Metabolomics

UCSF. Blood plasma samples were available for 91 subjects enrolled in the UCSF AD Research Centre. The ultra performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS) method used in this study has been previously published.[15] Twenty microliters of plasma per subject was required for analysis, with sample treatment being described elsewhere.[15,21] The method primarily detects lipids and has been shown to measure abundances of over 4500 metabolic features. The instruments included a Waters ACQUITY UPLC and Xevo Quadrupole Time-of-flight System (Waters, Milford, CT, USA). The Xevo Quadrupole Time-of-flight System was operated in both negative and positive ion modes. Samples were analyzed as one batch in a randomized order, with pooled plasma quality-control (QC) samples run between every 10 samples.

ADNI. Metabolite data were available for 853 blood serum samples. Twenty-four subjects had two samples included in the study. Targeted metabolomics analysis was performed using the AbsoluteIDQ p180 assay (Biocrates Life Sciences, Innsbruck, Austria) requiring 10 μl of serum per sample. The samples were run in 11 batches with two pooled QC samples present in each batch: one run before the samples and one afterward. More information on the assay, sample treatment and instruments can be downloaded from the ADNI website (adni.loni.usc.edu/).

Candidate protein assays (UCSF only)

The proteomics approach used in this study has been described elsewhere.[16] In short, a set of candidate proteins was quantified using single analyte sandwich enzyme-linked immunosorbent assays. In this study we investigated two proteins that have been replicated as NAB markers: FGG and PPY.[16,18,22,23]

NAB measurements

UCSF. Details of PET imaging are given elsewhere.[20] All PET scans used Pittsburgh compound B (11C-PiB) as the radioactive tracer. Scans were performed using two different scanner types, Biograph TruePoint 6 PET/computed tomography ($N = 9$) and Siemens ECAT EXACT HR PET ($N = 69$), and were processed using methods described by Lehmann *et al*.[24]

We considered two PET outcomes. Two experienced raters who were blinded to plasma and clinical data rated the scans as either high NAB or low NAB to give a dichotomous outcome. Second, the 50–70-min standardized uptake value ratio (SUVR) was used as a continuous outcome.[25]

ADNI. Details of PET imaging in ADNI (using both PiB and AV45 markers) and CSF measurements are detailed elsewhere (www.adni-info.org). PET end points were dichotomized into high and low NAB at the SUVR thresholds previously used in ADNI (1.5 for PiB and 1.11 for AV45). CSF measures of amyloid were taken from the data set 'UPENNBIOMK2' available on the ADNI website. The CSF measures were dichotomized at the previously published threshold (192 pg ml^{-1}). We combined the three amyloid end points into a combined amyloid end point to maximize sample size. A subject was classified as NAB-positive if at least one measurement indicated high brain amyloid burden, and classified as NAB-negative otherwise.

Statistical analysis

All statistical analyses were performed in R version 3.1.1.[26]

Data pre-processing

In UCSF, metabolic feature data were extracted from netCDF files using the R package 'XCMS'.[27] The package performed filtration and peak identification before matching peaks across samples and performing a retention time correction. Following data extraction, the negative- and positive-mode data were processed separately using the pipeline detailed in Supplementary Text 1. ADNI data were also processed using this pipeline. The processing included outlier removal, normalization through autoscaling and a log base 2 transformation as well as batch correction using the empirical Bayes method, ComBat.[28]

After pre-processing, the UCSF data collected in negative and positive modes were merged.

Protein data were subject to a natural logarithm transformation and screened for per sample, per protein outliers defined as values outside of 6 s.d.'s of the mean (as above). Each protein was autoscaled.

Single metabolic feature analysis

Single metabolic feature analysis was performed in UCSF for both NAB outcomes for each of the 2760 metabolic features detected. SUVR was linearly regressed against each metabolic feature in turn with *APOE* ε4 status and age included as covariates in the model. The *APOE* ε4 status is defined as 1 if a subject's genotype contained any ε4 alleles and 0 otherwise. Similarly, logistic regression was performed for the dichotomous outcome. In both cases, a Benjamini–Hochberg correction of the false discovery rate was applied.

Multiple metabolic feature analysis

Multiple metabolic feature analysis was performed on UCSF data using the R package 'caret'.[29] Partial least squares (PLS) and PLS discriminant analysis were used for the continuous and dichotomized outcomes, respectively. Ideally, we would have split the data into a training and test set; however, owing to relatively small sample size, this was not possible and a cross-validation (CV) approach was taken instead. All metabolic features, age and *APOE* ε4 status were included in the model building. The number of components to include was tuned using five-fold CV through the 'train' function. Recursive feature elimination was used to select a subset of variables using five-fold CV. The subset sizes considered varied from 2 to

99 in steps of 1 and from 100 to the total number of covariates in steps of 100. In PLS modeling, the lowest root mean squared error (RMSE) was used to select the best model, whereas for PLS discriminant analysis the highest accuracy was used. The function 'pickSizeTolerance' was then applied in an attempt to find a smaller subset of variables that maintained RMSE or accuracy to within 5% of the best model. We also built models using the 10 most important predictors, the maximum number of metabolic features we could feasibly identify. Model statistics resulting from five times CV within recursive feature elimination are presented in this report.

For comparison, we used five-fold CV to build a model based on age and *APOE* ε4 status alone using the 'train' function to tune the number of components as above. This method was used, despite the small number of predictors, to ensure continuity between modeling techniques. We checked that the results were consistent with those gained from a linear regression model. This model is referred to as the demographic-only model throughout.

Metabolic feature and protein joint analysis

The final multiple metabolic feature models were updated by adding proteins. Model building followed that of the demographic-only model detailed above. FGG and PPY were included both together and separately. We also modeled PPY and FGG (with and without age and *APOE* ε4 status) against continuous and dichotomized NAB without metabolic features for comparison.

Putative metabolite identification

Putative identification of selected metabolic features from statistical analysis was attempted using the median *m/z* and their corresponding retention time, initially using an in-house database and the Human Metabolome Database.[21,30] To enable the confirmation of features from the database-matching, fragmentation patterns were analyzed using level-two MS spectra.

Replication in ADNI

We searched the ADNI metabolite data for any of the metabolic features putatively identified in UCSF. Logistic regression models of the combined amyloid end point were built using individual metabolites as predictors, covarying for age and *APOE* ε4 status.

Code availability

All R codes used to generate this analysis are available from the corresponding author on request.

RESULTS

Data pre-processing

UCSF. The R package 'XCMS' extracted data for 248 metabolic features from negative ionization mode and 2807 metabolic features from positive ionization mode. We ran UPLC-MS/MS in the positive mode on 91 subject samples and 11 pooled QC samples. In the negative mode, data were available for 90 samples and 10 pooled QC samples.

ADNI. Data were available for 141 metabolites in 853 samples. This included 22 pooled QC samples and 24 replicates. As no documentation of technical replicates was given by ADNI, the first value was taken. This reduced the sample size to 829, including the 22 QC samples.

Figure 1 gives an overview of the pre-processing steps. In UCSF, this processing resulted in 78 subjects with dichotomous NAB and 76 subjects with continuous NAB. We had a total of 2760 metabolic features: 240 from the negative mode and 2520 from the positive mode. In ADNI, the processing resulted in 531 subjects with the combined amyloid end point and 116 metabolic features.

UCSF Positive mode	UCSF Negative mode	ADNI
Metabolic features identified		
Metabolic features (p) = 2807 Samples (N) = 91	Metabolic features (p) = 248 Samples (N) = 90	Metabolic features (p) = 141 Samples (N) = 831
UCSF: Remove any batches with only one sample ADNI:Remove duplicate samples		
p = 2807 N = 90	p = 248 N = 90	p = 141 N = 807
Remove metabolic features that elute before 1 minute or after 35 minutes		
p = 2520 N = 90	p = 241 N = 90	N/A
Identify per sample, per metabolic feature outliers and set to missing		
Remove metabolic features not present in at least 80% of samples		
p = 2520 N = 90	p = 240 N = 90	p = 117 N = 807
Remove samples with greater than 20% missingness		
p = 2520 N = 90	p = 240 N = 90	p = 117 N = 806
Investigate missingness Normalize using autoscaling Test normality of each metabolic feature Log base 2 transformation Adjust for batch effects using ComBat Use probabilistic principal component analysis to ensure QC samples cluster Impute data using 10 nearest neighbors		
Merge metabolic features discovered in positive and negative modes		
p = 2760 N = 89		N/A
Remove samples with missing APOE genotype		
p = 2760 N = 78		N/A
UCSF: Remove samples with missing continuous NAB from continuous NAB analysis ADNI: Remove samples with no amyloid measurement		
p = 2760 N = 76		p = 117 N = 531

Figure 1. Overview of pre-processing steps affecting the number of metabolic features and samples.

Cohort demographics

An overview of demographics for subjects included in the dichotomous NAB analysis is given in Table 1. The subjects used here have a wide range of diagnoses that can be grouped into four categories: AD, fronto-temporal dementia, MCI and healthy controls. Of these 78 subjects, 2 did not have SUVR available, reducing the number of subjects in the continuous analysis to 76. The demographics of this subpopulation are given in Supplementary Table 1. It is important to note that the population is relatively balanced in terms of age and scanner type between high and low NAB groups.

An overview of demographics for subjects included in the ADNI replication analysis is also shown in Table 1.

Single metabolic feature analysis

For both continuous and dichotomized NAB, no metabolic features passed a *q*-value threshold of 0.1. Supplementary Tables 2 and 3 give full results.

Multiple metabolic feature analysis

Continuous NAB. The multiple metabolite model with the lowest error was found for 100 predictors, all of which were metabolic features (CV RMSE = 0.53, CV $R^2 = 0.10$). A tolerance set was generated to maintain error (that is, CV RMSE) within 5% of the

Table 1. Cohort demographics

UCSF	*Total* (N = *78*)	*Low NAB* (N = *48*)	*High NAB* (N = *30*)	P-*value*
Median NAB SUVR (IQR)[a]	1.3 (0.9)	1.2 (0.1)	2.3 (0.4)	—
Plasma sample median days in storage	1354.5 (560.8)	1400 (432.8)	1247 (835.3)	0.435
Median number of days' difference between sample collection and scan (IQR)	18.5 (69.8)	18.5 (62.3)	16 (101.5)	0.963
Median age (IQR)	65.5 (10.7)	65.8 (9.1)	63.1 (12.7)	0.472
Median MMSE (IQR)	25.5 (6.0)	27 (4.3)	22.5 (9.8)	< 0.001
Scanner type (%)				
Biograph	9 (11.5)	6 (12.5)	3 (10.0)	> 0.999
Siemens	69 (88.5)	42 (87.5)	27 (90.0)	
Gender (%)				
Female	32 (41.0)	18 (37.5)	14 (46.7)	0.482
Male	46 (59.0)	30 (62.5)	16 (53.3)	
APOE ε4 status (%)				
0	57 (73.1)	38 (81.2)	18 (60.0)	0.065
1	21 (26.9)	10 (18.8)	12 (40.0)	
Diagnosis (%)				
AD	24 (30.8)	2 (4.2)	22 (73.3)	< 0.001
FTD	48 (61.5)	42 (87.5)	6 (20.0)	
HC	4 (5.1)	3 (6.3)	1 (3.3)	
MCI	2 (2.6)	1 (2.1)	1 (3.3)	
ADNI	N = *531*	N = *265*	N = *266*	
Median age (IQR)	75.1 (8.70)	75.8 (8.60)	74.3 (8.78)	0.497
Gender (%)				
Female	213 (40.1)	110 (41.5)	103 (38.7)	0.536
Male	318 (59.9)	155 (58.5)	163 (61.3)	
Median years in education (IQR)	16 (4)	16 (4)	16 (4)	0.505
APOE ε4 status (%)				
0	279 (52.5)	115 (43.4)	164 (61.7)	< 0.001
1	252 (47.5)	150 (56.6)	102 (38.3)	
Median MMSE (IQR)	27 (5)	26 (6.75)	28 (4)	0.001
Diagnosis (%)				
Other	88 (16.6)	36 (13.6)	52 (19.5)	< 0.001
Dementia	157 (29.5)	101 (38.1)	56 (21.0)	
MCI	172 (32.5)	80 (30.2)	92 (34.6)	
HC	114 (21.5)	48 (18.1)	66 (24.8)	

Abbreviations: AD, Alzheimer's disease; ADNI, Alzheimer's disease neuroimaging initiative; FTD, fronto-temporal dementia; HC, healthy control; IQR, interquartile range; MCI, mild cognitive impairment; MMSE, mini mental state exam; NAB, neocortical amyloid burden; SUVR, standardized uptake value ratio; UCSF, University of California, San Francisco. [a]This is based on those subjects with SUVR available ($N = 76$; low NAB $N = 48$; high NAB $N = 28$). Kruskal–Wallis X^2 was used to test between high and low groups for continuous demographic variables. Fisher's exact was used to test between high and low groups for categorical demographic variables.

value achieved by the optimal model (0.53). The reduced model contained 17 of these metabolic features (CV RMSE = 0.55, CV $R^2 = 0.07$).

The 10 predictor models contained only one component (CV RMSE = 0.56, CV $R^2 = 0.05$). Cross-validated model statistics are given in Table 2, illustrating that the models including metabolic features do not outperform age and *APOE* in this training data. For information on the metabolic features included in the final models see Table 3.

Addition of the proteins FGG and PPY to the 17-metabolic-feature model increased cross-validated R^2 to 0.57, explaining more variation in NAB than metabolic features or proteins alone ($R^2 = 0.07$ and 0.21, respectively).

Dichotomized NAB. The best model and tolerance set model were the same, both containing five-metabolic-feature predictors (CV accuracy = 0.72, CV sensitivity = 0.65, CV specificity = 0.76). Model statistics for the final models are given in Table 2. We see an improved accuracy of 72% compared with age and *APOE* alone at 58%. For information on the five metabolic features included in the final model see Table 3.

The addition of the protein FGG to the five-metabolic-feature model increased accuracy to 79%, with sensitivity and specificity both above 70% (71% and 84%, respectively). The two protein models (FGG and PPY only) gave an identical accuracy to the five-metabolic-feature model at 72%, driven by a high specificity (93%).

Putative metabolite identification. We aimed to putatively identify the five metabolic features that were included in the final model of dichotomized NAB (Figure 2). We were able to identify four of these five metabolic features. No suitable surrogate metabolic feature was available for the unidentified metabolite.

Table 2. Multiple metabolic feature analysis

Continuous NAB models	R^2	*RMSE*	
Tolerance set (17 metabolic features)	0.07	0.55	
10 Metabolic features	0.05	0.56	
Age and *APOE* status	0.12	0.55	
Tolerance set (17 metabolic features) with FGG	0.57	0.37	
Tolerance set (17 metabolic features) with PPY	0.57	0.38	
Tolerance set (17 metabolic features) with FGG and PPY	0.57	0.37	
FGG and PPY	0.21	0.49	
FGG with *APOE* status and age	0.09	0.53	
PPY with *APOE* status and age	0.01	0.54	
FGG and PPY with *APOE* status and age	0.08	0.52	
Dichotomized NAB models	*Accuracy*	*Sensitivity*	*Specificity*
Five metabolic features	0.72	0.65	0.76
Age and *APOE* status	0.58	0.10	0.88
Tolerance set (five metabolic features) with FGG	0.79	0.71	0.84
Tolerance set (five metabolic features) with PPY	0.75	0.58	0.86
Tolerance set (five metabolic features) with FGG and PPY	0.78	0.65	0.89
FGG and PPY	0.72	0.43	0.93
FGG with *APOE* status and age	0.70	0.43	0.90
PPY with *APOE* status and age	0.58	0.26	0.83
FGG and PPY with *APOE* status and age	0.65	0.39	0.86

Abbreviations: FGG, fibrinogen gamma chain; NAB, neocortical amyloid burden; PPY, pancreatic polypeptide; RMSE, root mean square error. Table shows cross-validated model statistics for continuous NAB.

Table 3. Metabolic features included in the multiple metabolic feature models

Continuous NAB tolerance set model			*Continuous NAB 10 predictor model*			*Dichotomized NAB model*		
Mode	*Median* m/z	*Median retention time (min)*	*Mode*	*Median* m/z	*Median retention time (min)*	*Mode*	*Median* m/z	*Median retention time (min)*
Positive	184.10	2.85	Positive	184.10	2.85	Positive[a]	647.59	10.69
Positive	370.41	11.51	Positive	370.41	11.51	Positive[a]	648.59	10.69
Positive	565.64	18.24	Positive	565.64	18.24	Negative[a]	775.68	16.38
Positive	700.62	17.09	Positive	700.62	17.09	Positive[a]	778.63	14.94
Positive	718.65	17.20	Positive	718.65	17.20	Negative	829.66	16.52
Negative	726.62	18.54	Negative	774.62	18.38			
Positive	755.64	13.93	Negative[a]	775.68	16.38			
Negative	774.62	18.38	Negative	775.63	18.38			
Negative[a]	775.68	16.38	Positive	776.66	18.53			
Negative	775.63	18.38	Negative	829.66	16.52			
Positive	776.66	18.53						
Positive[a]	778.63	14.94						
Positive	784.68	16.19						
Positive	791.68	16.85						
Negative	829.66	16.52						
Positive	903.81	21.38						
Positive	903.86	29.42						

Abbreviation: NAB, neocortical amyloid burden. [a]Identified metabolic feature.

One of the four metabolic features was discovered in negative-mode UPLC-MS/MS (median $m/z = 775.68$) and has been identified as a phosphatidylethanolamine (PE 39:7). The remaining metabolic features were discovered in the positive mode. The metabolic feature with median $m/z = 647.59$ and an isotope (median $m/z = 648.59$) are likely to be anandamide (linoleoyl ethanolamide (2M+H)).[31] As expected, these isotopes are highly correlated (Pearson's correlation coefficient = 0.966). Fragmentation patterns of the metabolic feature with the median $m/z = 778.63$ suggest a phosphatidylcholine (PCaa 36:6).

Replication in ADNI. One of the four putatively identified metabolites from UCSF was found in the ADNI data: PCaa 36:6. In the logistic regression model of the combined amyloid end point, PCaa 36:6 had an estimate of −0.729 ($P = 0.066$).

DISCUSSION

To the best of our knowledge, this is the first study to investigate associations between blood metabolites and amyloid burden in the brain. We have used non-targeted metabolomics to predict NAB in subjects from the UCSF AD research center. We also present the first analysis to combine protein and metabolite data in the search for a biomarker for AD.

We found a panel of five metabolic features that predicted amyloid positivity with an accuracy of 72%. If the model specificity (76%) seen here is maintained in a replication study, it could be

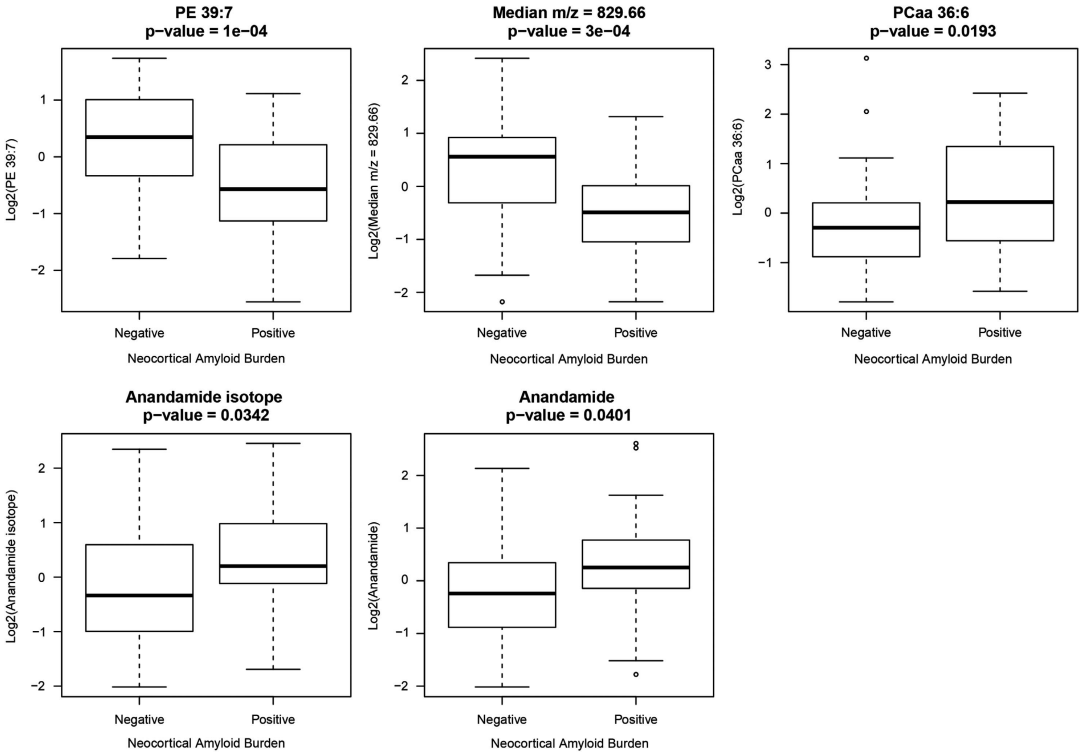


Figure 2. Boxplots showing metabolic feature levels between high and low neocortical amyloid burden (NAB) groups for the five metabolic features included in the final model of dichotomized NAB. Student's *t*-test was used to generate a *P*-value.

useful in a screening setting where a large proportion of subjects would have high amyloid burden. As the metabolite panel correctly identified subjects with low amyloid levels, 76% of the time it could be useful in reducing the number of patients with low amyloid burden unnecessarily subjected to further procedures. Interestingly, no metabolic feature model retained age or *APOE* ε4 status. This population appeared relatively balanced with respect to these two variables, possibly accounting for the lack of inclusion. Alternatively, effects of age and *APOE*, which are well known to be associated with amyloid burden, could be accounted for in surrogate metabolic feature variables.[6,7] Analysis of single metabolic features gave no significant results. However, low statistical power in the current study means that this approach should not be ruled out in further study of larger cohorts.

We were able to putatively identify four of the five metabolic features included in the model of dichotomized NAB. Of those identified, one was a phosphatidylcholine compound (PCaa 36:6). PCs are a group of compounds previously implicated in AD by Whiley *et al.* and others.[12,15] In particular, PCaa 36:6 was included in the 10-lipid panel suggested by Mapstone *et al.* to predict conversion to amnestic MCI or AD with 90% accuracy. The association is in the opposite direction to that seen here, which could be explained by differences in disease stage between the cohorts. PCs are phospholipids that form a substantial component of biological membranes, and in this study show increased abundance in subjects with high NAB. Chung *et al.* state that PCs improve memory in mouse models, corresponding with the direction of association seen by Mapstone *et al.*[13] However, a Cochrane review has surmised that there is not sufficient evidence to extend this conclusion to humans.[33]

We were able to test associations of PCaa 36:6 in the ADNI cohort. We saw a direction of association concurrent with that seen by Mapstone *et al.* but opposite to that seen in the UCSF cohort.[32] Subjects in the ADNI cohort are diagnostically more similar to those used by Mapstone *et al.*, which could account for this similarity. Further, as we could only test the one metabolite, it is possible that this discrepancy is because we could not include the other four metabolites. ADNI is currently the only other cohort that has both metabolite and amyloid data available, and consequently this is the maximum extent of replication we can perform. It is essential that further attempts at replication are made in larger, independent studies.

We were also able to identify a PE (39:7). PEs are also a subtype of phospholipids that can be found in biological membranes. Interestingly, in humans they are largely found in tissues of the central nervous system and when methylated yield phosphatidylcholines.[34] PEs are also implicated in prion disease, where they cause aggregation of the prion protein.[35] In this study PE 39:7 was reduced in subjects with high NAB. PEs are also substrates for the synthesis of the final metabolic feature we were able to identify: anandamide.[36] Anandamide is an endogenous cannabinoid neurotransmitter that, on connection with receptors in the cell membrane, reduces the release of other neurotransmitters in the brain.[37] Anandamide is fat soluble, allowing it to pass through the blood–brain barrier and is made in areas of the brain important in memory. It is hypothesized that anandamide is involved in the creation and deletion of short-term connections between nerve cells.[38] In support of this theory, the presence of anandamide has been shown to impair memory in rats.[38] In this study anandamide is reduced in subjects with high NAB. This

supports findings by Jung *et al.*[39] who see an Aβ-dependent association of anandamide with cognitive decline in samples of brain tissue.

This study shows for the first time that the addition of candidate proteins (FGG and PPY) to metabolic feature models improves results. These results are promising and warrant further study while reinforcing the idea that a multi-modal approach may be more effective in AD biomarker discovery than single modality approaches.

Although the results we present here are interesting, and we are reassured by the fact that the findings make biological sense, this study does have limitations—in particular, a lack of test data and the difference of direction of association for PCaa 36:6 in ADNI. Without a full independent test set it is likely that model statistics will be inflated, and therefore the results should be interpreted cautiously. Our preference would have been to split the data into a training and test set; however, the relatively small sample size made this suggestion infeasible. Instead, we choose to use a five-fold CV approach in this study. It is essential to validate this work in independent cohorts of a larger size, for example, in an asymptomatic cohort with high amyloid levels, to reflect the populations eligible for trials such as the A4 trial.[3] A further issue caused by small sample size is a lack of statistical power. This could be causing the substantial differences in R^2 seen in the continuous NAB analysis and provides further rationale for this work to be replicated in larger cohorts.

A further limitation of this study is the confounding factor of diagnosis: the majority of subjects with high NAB have AD, whereas the majority of subjects with low NAB are diagnosed with fronto-temporal dementia. It is therefore impossible to tell whether the markers we identify here differentiate between high and low NAB or AD and fronto-temporal dementia. Both applications are important and interesting; however, it is vital that we aim to understand this confounding in future studies perhaps through similar analysis in an AD-only cohort.

In further research, targeted metabolite analysis would be beneficial. With an increased annotation, the biological understanding of any findings would grow, potentially deepening our knowledge of the disease mechanism of AD. Metabolite identification using the current methods is time-consuming, often inconclusive and can only be confirmed when the pure standard compounds are available. Further, the presence of annotated data would enable pathway analysis and more ready replication of findings. The data available in ADNI begin to work toward this.

CONCLUSION

This study used metabolomic information to predict NAB in subjects from the UCSF AD Research Centre. Five metabolic features identified subjects with high NAB with 72% accuracy. We were able to identify four metabolic features from this panel (PCaa 36:6, PE 39:7 and Anandamide and an isotope) and improve the model further with the addition of FGG protein measures (accuracy = 79%). If replicated in large, independent studies, these metabolic features and proteins could form the basis of a blood test with potential for enrichment of amyloid pathology in anti-amyloid trials.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

We thank the UCSF Memory and Aging Centre, all the participants who took part in this study and the clinicians who referred participants. The UCSF study was funded by the US National Institutes of Health grants K23-AG031861, R01-AG027859, R01-AG032306, R01- AG038791, P01-AG1972403 and P50- AG023501; the State of California Department of Health Services; the Alzheimer's Association; and the John Douglas French Alzheimer's Foundation. ADNI data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI; National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica; Biogen; Bristol-Myers Squibb Company; CereSpir; Eisai; Elan Pharmaceuticals; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche and its affiliated company Genentech; Fujirebio; GE Healthcare; IXICO; Janssen Alzheimer Immunotherapy Research & Development; Johnson & Johnson Pharmaceutical Research & Development; Lumosity; Lundbeck; Merck & Co; Meso Scale Diagnostics; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. This work represents independent research part funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement number 115372, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution. SJK is supported by an MRC Career Development Award in Biostatistics (MR/L011859/1). NV is funded by the Alzheimer's Society. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

REFERENCES

1 Corbett A, Ballard C. New and emerging treatments for Alzheimer's disease. *Exp Opin Emerg Drugs* 2012; **17**: 147–156.
2 Salloway S, Sperling R, Fox N, Blennow K, Klunk W, Raskind M *et al.* Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. *N Engl J Med* 2014; **370**: 322–333.
3 Sperling R, Rentz D, Johnson K, Karlawish J, Donohue M, Salmon D *et al.* The A4 study: stopping AD before symptoms begin? *Sci Transl Med* 2014; **6**: 228fs13.
4 Villemagne V, Pike K, Chételat G, Ellis K, Mulligan R, Bourgeat P *et al.* Longitudinal assessment of Aβ and cognition in aging and Alzheimer disease. *Ann Neurol* 2011; **69**: 181–192.
5 Jack C, Knopman D, Jagust W, Petersen R, Weiner M, Aisen P *et al.* Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *Lancet Neurol* 2013; **12**: 207–216.
6 Jansen W, Ossenkoppele R, Knol D, Tijms B, Scheltens P, Verhey F *et al.* Prevalence of cerebral amyloid pathology in persons without dementia: a meta-analysis. *JAMA* 2015; **313**: 1924–1938.
7 Ossenkoppele R, Jansen W, Rabinovici G, Knol D, van der Flier W, van Berckel B *et al.* Prevalence of amyloid pet positivity in dementia syndromes: a meta-analysis. *JAMA* 2015; **313**: 1939–1949.
8 Kaddurah-Daouk R, Zhu H, Sharma S, Bogdanov M, Rozen S, Matson W *et al.* Alterations in metabolic pathways and networks in Alzheimer's disease. *Transl Psychiatry* 2013; **3**: e244–e248.
9 Bazenet C, Lovestone S. Plasma biomarkers for Alzheimer's disease: much needed but tough to find. *Biomarkers Med* 2012; **6**: 441–454.
10 Montagne A, Barnes SR, Sweeney MD, Halliday MR, Sagare AP, Zhao Z *et al.* Blood-brain barrier breakdown in the aging human hippocampus. *Neuron* 2015; **85**: 296–302.
11 Ellis B, Hye A, Snowden S. Metabolic modifications in human biofluids suggest the involvement of sphingolipid, antioxidant, and glutamate metabolism in Alzheimer's disease pathogenesis. *J Alzheimer's Dis* 2015; **46**: 313–327.
12 Orešič M, Hyötyläinen T, Herukka S, Sysi-Aho M, Mattila I, Seppänan-Laakso T *et al.* Metabolome in progression to Alzheimer's disease. *Transl Psychiatry* 2011; **1**: e57–e59.
13 Mapstone M, Cheema A, Fiandaca M, Zhong X, Mhyre T, MacArthur L *et al.* Plasma phospholipids identify antecedent memory impairment in older adults. *Nat Med* 2014; **20**: 415–420.
14 Proitsi P, Kim M, Whiley L, Pritchard M, Leung R, Soininen H *et al.* Plasma lipidomics analysis finds long chain cholesteryl esters to be associated with Alzheimer's disease. *Transl Psychiatry* 2015; **5**: e494.

15 Whiley L, Sen A, Heaton J, Proitsi P, García-Gómez D, Leung R *et al.* Evidence of altered phosphatidylcholine metabolism in Alzheimer's disease. *Neurobiol Aging* 2014; **35**: 271–278.
16 Ashton N, Kiddle S, Graf J, Ward M, Baird A, Hye A *et al.* Blood protein predictors of brain amyloid for enrichment in clinical trials? *Alzheim Dement* 2015; **1**: 48–60.
17 Sabuncu MR, Buckner RL, Smoller JW, Lee PH, Fischl B, Sperling RA *et al.* The association between a polygenic Alzheimer score and cortical thickness in clinically normal subjects. *Cereb Cortex* 2012; **22**: 2653–2661.
18 Voyle N, Baker D, Burnham S, Covin A, Zhang Z, Sangurdekar D *et al.* Blood protein markers of neocortical amyloid-β burden: a candidate study Using SOMAscan technology. *J Alzheim Dis* 2015; **46**: 947–962.
19 Bettcher B, Watson C, Walsh C, Lobach I, Neuhaus J, Miller J *et al.* Interleukin-6, age, and corpus callosum integrity. *PLoS ONE* 2014; **9**: e106521.
20 Rabinovici G, Rosen H, Alkalay A, Kornak J, Furst A, Agarwal N *et al.* Amyloid vs FDG-PET in the differential diagnosis of AD and FTLD. *Neurology* 2011; **77**: 2034–2042.
21 Whiley L, Godzien J, Ruperez F, Legido-Quigley C, Barbas C. In-vial dual extraction for direct LC-MS analysis of plasma for comprehensive and highly reproducible metabolic fingerprinting. *Anal Chem* 2012; **84**: 5992–5999.
22 Kiddle S, Thambisetty M, Simmons A, Riddoch-Contreras J, Hye A, Westman E *et al.* Plasma based markers of [11C] PiB-PET brain amyloid burden. *PLoS ONE* 2012; **7**: e44260.
23 Burnham S, Faux N, Wilson W, Laws S, Ames D, Bedo J *et al.* A blood-based predictor for neocortical Aβ burden in Alzheimer's disease: results from the AIBL study. *Mol Psychiatry* 2014; **19**: 519–526.
24 Lehmann M, Ghosh P, Madison C, Laforce R, Corbetta-Rastelli C, Weiner M *et al.* Diverging patterns of amyloid deposition and hypometabolism in clinical variants of probable Alzheimer's disease. *Brain* 2013; **136**: 844–858.
25 Lopresti B, Klunk W, Mathis C, Hoge J, Ziolko S, Lu X *et al.* Simplified quantification of pittsburgh compound B amyloid imaging PET studies: a comparative analysis. *J Nucl Med* 2005; **46**: 1959–1972.
26 Core Team R. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria, 2014.
27 Smith C, Want E, O'Maille G, Abagyan R, Siuzdak G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal Chem* 2006; **78**: 779–787.
28 Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007; **8**: 118–127.
29 Kuhn JM, with contributions from Wing, Weston S, Williams A, Keefer C, Engelhardt A, Cooper T *et al.* caret: Classification and Regression Training, 2014. R package version 6.0-35.
30 Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, Liu Y *et al.* HMDB 3.0–the human metabolome database in 2013. *Nucleic Acids Res* 2013; **41**: D801–D807.
31 Schreiber D, Harlfinger S, Nolden B, Gerth C, Jaehde U, Schömig E *et al.* Determination of anandamide and other fatty acyl ethanolamides in human serum by electrospray tandem mass spectrometry. *Anal Biochem* 2007; **361**: 162–168.
32 Chung S, Moriyama T, Uezu E, Uezu K, Hirata R, Yohena N *et al.* Administration of phosphatidylcholine increases brain acetylcholine concentration and improves memory in mice with dementia. *J Nutr* 1995; **125**: 1484–1489.
33 Higgins J, Flicker L. Lecithin for dementia and cognitive impairment. *Cochrane Database Syst Rev* 2000; **4**, Art. No. CD001015.
34 Vance J, Tasseva G. Formation and function of phosphatidylserine and phosphatidylethanolamine in mammalian cells. *Biochim Biophys Acta* 2013; **1831**: 543–554.
35 Deleault N, Piro J, Walsh D, Wang F, Ma J, Geoghegan J *et al.* Isolation of phosphatidylethanolamine as a solitary cofactor for prion formation in the absence of nucleic acids. *Proc Natl Acad Sci USA* 2012; **109**: 8546–8551.
36 Jin X, Okamoto Y, Morishita J, Tsuboi K, Tonai T, Ueda N. Discovery and characterization of a Ca2+ independent phosphatidylethanolamine N-acyltransferase generating the anandamide precursor and its congeners. *J Biol Chem* 2007; **282**: 3614–3623.
37 Stella N, Schweitzer P, Piomelli D. A second endogenous cannabinoid that modulates long-term potentiation. *Nature* 1997; **388**: 773–777.
38 Mallet P, Beninger R. The endogenous cannabinoid receptor agonist anandamide impairs memory in rats. *Behav Pharmacol* 1996; **7**: 276–284.
39 Jung KM, Astarita G, Yasar S, Vasilevko V, Cribbs DH, Head E *et al.* An amyloid β_{42} – dependent deficit in anandamide mobilization is associated with cognitive dysfunction in Alzheimer's disease. *Neurobiol Aging* 2012; **33**: 1522–1532.

Supplementary Information accompanies the paper on the Translational Psychiatry website (http://www.nature.com/tp)

Chapter 5

Do peripheral markers help to predict amyloid burden in a non-demented population? A Bayesian approach

Does genetic risk help to predict amyloid burden in a non-demented population? A Bayesian approach.

Nicola Voyle[a], Willemijn Jansen[b], Aoife Keohane[a], Hamel Patel[a], Amos Folarin[a], Stephen Newhouse[a], Caroline Johnston[a], Kuang Lin[c], Pieter Jelle Visser[d,e], Angela Hodges[c], Richard JB Dobson * †[a,f,g], Steven J Kiddle * †[a,h], for the Alzheimer's Disease Neuroimaging Initiative ‡[1], and EDAR and DESCRIPA study groups[1]

[a]MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

[b]Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Alzheimer Center Limburg, Maastricht University, The Netherlands

[c]Department of Old Age Psychiatry, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

[d]Department of Neurology and Alzheimer Center, VU University Medical Center, Amsterdam, The Netherlands

[e]Helen Wills Neuroscience Institute, University of California, Berkeley

[f]NIHR Biomedical Research Centre for Mental Health and Biomedical Research Unit for Dementia at South London and Maudsley NHS Foundation, London, UK

[g]Farr Institute of Health Informatics Research, UCL Institute of Health Informatics, University College London, London, UK

[h]MRC Biostatistics Unit, Cambridge Biomedical Campus, Cambridge Institute of Public Health, Forvie Site, Robinson Way, Cambridge, UK.

*Joint senior author

†Corresponding authors: SGDP Centre, IoPPN, 16 De Crespigny Park, Denmark Hill, London, SE5 8AF; richard.j.dobson@kcl.ac.uk; steven.kiddle@kcl.ac.uk; Telephone 020 7848 0924; Fax 020 7848 0866

‡Data used in preparation of this article were obtained from the Alzheimers Disease Neuroimaging Initiative (ADNI) database

(adni.loni.usc.edu). The investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data, but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at http //adni.loni.usc.edu/wp−content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

Abstract

INTRODUCTION: In this study we investigate the association between Aβ levels in cerebrospinal fluid (CSF) and genetic risk in a non-demented population. This paper presents the first analysis to use a Bayesian methodology in this area.

METHODS: Data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the EDAR* and DESCRIPA** studies was used in a Bayesian logistic regression analysis. We modeled CSF Aβ burden using age, diagnosis (healthy control or mild cognitive impairment), *APOE* and a polygenic risk score (PGRS) associated with Alzheimer's Disease (AD). We compared models built using informative priors on age, diagnosis and *APOE* with non-informative priors on all variables.

RESULTS: The use of informative priors did not improve model performance in the majority of cases. Models using only age, diagnosis and *APOE* genotype showed the best predictive ability.

DISCUSSION: A previous study indicated that a PGRS of AD case/control status was associated with CSF Aβ burden in healthy controls. The current study suggests that this association does not lead to models that are more predictive of amyloid positivity than already known factors such as age and *APOE*.

* 'Beta amyloid oligomers in the early diagnosis of AD and as marker for treatment response'

** 'Development of screening guidelines and criteria for pre-dementia Alzheimers disease'

KEYWORDS: Amyloid; Alzheimer's Disease; Multi-modal; Polygenic Risk Score; Bayesian; Gene Expression Risk Score; Blood; Biomarker; Protein; Metabolite.

1 Introduction

It is hypothesized that late onset AD is caused by the presence of Aβ plaques in the brain tissue and hyperphosphorylated tau tangles in the neurons [1]. Hypothetical models and longitudinal studies indicate that Aβ pathology begins to develop up to 20 years prior to symptomatic changes [2, 3, 4]. This provides a window of opportunity for disease-modifying treatments, provided accurate and sensitive diagnostic tools are available. Existing tools for identifying the presence of pathology include measurements of amyloid and tau in CSF or in brain by the use of positron emission tomography (PET) imaging. Although these tests are reasonably accurate they are invasive and expensive. Peripheral biomarkers are sought as an intermediate step to provide cost-effective enrichment of people with high risk profiles for clinical trials, particularly secondary prevention trials of anti-amyloid therapeutics. In this study we aimed to address this by investigating the degree to which a genetic risk score was predictive of CSF Aβ in a non-demented population.

AD has been identified as a complex disease meaning its presence or absence is determined by environmental and genetic risk factors. Heritability is estimated to be 50-70% [5, 6, 7]. Over 20 risk loci have been identified for AD with Lambert *et al.* providing the most comprehensive Genome Wide Association Study (GWAS) to date [8, 9, 10, 11, 7]. However, it is estimated that currently identified SNPs can only explain 16-33% of phenotypic variation [9, 10, 7]. GWAS of Aβ endpoints are also highlighting promising results but sample sizes are considerably smaller so validation in larger numbers is needed [12]. An approach to consolidating the combined effect of several smaller genetic contributions to disease is found in polygenic risk scores (PGRS). A PGRS is calculated as a weighted sum and represents the cumulative effect of these smaller genetic effects . They have been shown to be informative in understanding genetic contributions to phenotypes beyond AD clinical diagnosis. For example, Sabuncu *et al.* found an AD case/control PGRS associated with Aβ burden in healthy controls ($p < 0.0001$). It has not been assessed whether this would provide predictive utility for enriching populations for prevention trials [13, 14, 15].

It is well known that model estimates, and often predictive ability, become more reliable as the number of individuals included in an analysis increases. However, in AD research the amalgamation of studies to create large datasets is often unfeasible due to differences in study populations and data collection methods. Initiatives such as the European Medical Information Framework (EMIF) are aiming to rectify this, but large populations with both multi-modal biomarker data and amyloid pathology measures are not yet available (www.emif.eu).

However, there is substantial information available on the associations between demographic variables and prevalence of Aβ burden, as discussed in meta-analyses by Jansen *et al.* and Ossenkoppele *et al.* [4, 16]. The former of these studies concentrated on persons without dementia (N=7583). The study concluded that age, *APOE* genotype and presence of cognitive impairment were associated with Aβ burden. No equivalent studies are available for tau pathology and consequently this work focuses on amyloid alone.

The present study uses data from the ADNI, EDAR and DESCRIPA cohorts, to investigate genetic risk as a blood biomarker of Aβ burden using a Bayesian methodology [17]. This study includes older individuals who do not have a clinical diagnosis of AD and are at a variable risk of developing the disease. This is the first study to use a Bayesian framework in AD blood biomarker research with the aim of investigating genetic risk as a marker in blood that could support strategies for identifying individuals at high risk of developing disease for recruitment into clinical trials. The models created combine age, diagnosis (control or MCI) and *APOE* genotype with a PGRS. A previous study has shown a case/control PGRS associated with Aβ in healthy controls. We aim to investigate whether this association translates to predictive ability. We hypothesized that by informing estimates for demographic variables using the large meta-analysis presented by Jansen *et al.* we would create more robust models. The Bayesian methodology used here has been made accessible to future researchers though the development of a simple graphical user interface (GUI).

2 Methods

2.1 Cohorts

2.1.1 EDAR

EDAR is a prospective, longitudinal study with centers at multiple European sites. The study aims to examine and evaluate biomarkers of early AD and treatment response [17]. For more information see www.edarstudy.eu. Our access to samples and clinical and phenotypic information from the EDAR study was enabled by EMIF.

2.1.2 DESCRIPA

DESCRIPA is a prospective, multi-center study based in Europe and coordinated by the European AD Consortium. The study focused on collecting data from non-demented subjects with the aim of developing

screening guidelines and clinical criteria for AD in non-demented subjects. Further details of this study can be found in Visser *et al.* [18].

2.1.3 ADNI

ADNI is a longitudinal cohort study aiming to validate the use of biomarkers in AD clinical trials and diagnosis. Data used in the preparation of this article were obtained from the ADNI database (adni.loni.usc.edu). The ADNI study was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether biological markers and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and AD. For information, see www.adni-info.org. ADNI was approved by the institutional review boards of all participating institutions, and written informed consent was obtained for all participants.

The ADNI study comprises three stages. ADNI 1 is the initial study (target N = 800). ADNI GO contains a subset of the controls and MCI participants from ADNI 1 and is supplemented by additional individuals with MCI (target N = 700). ADNI 2 enhances ADNI 1 and ADNI GO further with the inclusion of new participants in all diagnostic groups (target N = 1350).

This study uses data from ADNI 1 and the ADNI 2 and ADNI GO sub-groups, referred to as ADNI 2 from here onwards.

2.2 Genetics

Samples from EDAR and DESCRIPA were genotyped on the Illumina HumanOmniExpressExome-8v1.2 BeadChip and processed together (N = 336) [19]. This BeadChip contains 960,919 markers of which 273,000 represent functional exomic markers. This is the first publication to present genotype information for EDAR and DESCRIPA. The data was processed in Genome Studio (as described at bit.ly/1VpRclH) before being run through the rare variant caller Zcall [20] (as described here: bit.ly/1YKHYhK). ADNI 1 samples were run on the Human610-Quad BeadChip (N = 818) while ADNI 2 and ADNI GO samples were run on the Illumina HumanOmniExpress BeadChip (N = 432). The HumanOmniExpress BeadChip is similar to that used in the EDAR and DESCRIPA studies but does not include exomic markers, while the Human610-Quad BeadChip is older. Details of the genotyping protocols followed in ADNI are given elsewhere [21].

The cohorts were subject to quality control and imputation separately, as described in Coleman *et al.*

[22]. In short, the data was filtered to ensure a minor allele frequency of greater than 5% for all SNPs before removal of rare variants and subjects with high levels of missing data. SNPs that differed significantly ($p < 0.00001$) from the Hardy-Weinberg equilibrium were removed. The data was pruned for SNPs in linkage disequilibrium and for genetically similar individuals. Finally, the data was imputed using reference files from the 1000 Genomes Project. [23]

2.3 Amyloid measurements

Throughout this study amyloid measurements in cerebrospinal fluid (CSF) are used as the endpoint of interest. All Aβ measurements were dichotomized as detailed in the meta-analysis from Jansen *et al.*. The distribution of amyloid burden in all studies was bimodal (as expected) making this dichotomization biologically relevant. Low CSF Aβ is referred to as 'abnormal' Aβ burden while high CSF Aβ is referred to as 'normal'. The details for each study are as follows:

2.3.1 EDAR

Details of CSF collection and analysis can be found at www.edarstudy.eu. In brief, CSF measurements were collected using the Alzbio3 Luminex assay in one batch at the end of the study. CSF amyloid measurements were dichotomized at the previously published threshold of 389pg/ml.

2.3.2 DESCRIPA

Details of CSF measurements in DESCRIPA have been described elsewhere [19]. In brief, CSF measurements were analyzed in one laboratory and collected using single-parameter ELISA kits (Innogenetics, Ghent, Belgium). CSF amyloid samples were dichotomized using the previously published threshold of 550pg/ml.

2.3.3 ADNI

For ADNI, datasets used to extract CSF measures of amyloid were chosen to maximize sample size. The dataset 'UPENNBIOMK2' was used for ADNI 1 and 'UPENNBIOMK6' for ADNI 2 and ADNI GO. Both datasets contain CSF measurements collected using the xMAP Luminex platform and Innogenetics immunoassay kits. The CSF measures were dichotomized at the previously published threshold (192 pg/ml).

2.4 Statistical analysis

All statistical analysis was performed in R version 3.1.1 [24]. Models were built in ADNI 1 data and tested in data from EDAR, DESCRIPA and ADNI 2. We built models including age, diagnosis (healthy control or MCI) and *APOE* genotype as covariates ('basic model') and models including these variables with the addition of a PGRS ('PGRS model'). The predictive ability of each model was quantified using accuracy, sensitivity, specificity and area under the Receiver Operating Characteristic (ROC) curve [25, 26, 27].

2.4.1 PGRS

PGRS were created using the software package PRSice [28]. Effect sizes from stage 1 of the International Genomics of Alzheimer's Project (IGAP) case/control GWAS were used as the weights to generate the risk score (N = 54,162, number of SNPs = 7,055,881) [8]. We used 0.5 as the p-value threshold for inclusion in the PGRS. This threshold showed the most significant association with case/control diagnosis in the large IGAP PGRS study [15]. The genetic region coding for *APOE* was removed from all scores and included as a covariate in modeling.

2.4.2 Data analysis

This study aimed to predict dichotomized amyloid burden using genetic risk in a Bayesian logistic regression. The method was implemented using the R function 'MCMClogit' in the 'MCMCpack' package [29]. Models were built using a Metropolis sampler with 100,000 MCMC iterations, with the first 3,000 discarded as burn-in. This number of iterations ensured the ratio of standard deviation to Monte Carlo Standard Error (MCSE) was less than 5% for all parameters.

The variables included in the 'basic model' were chosen based on the meta-analysis published by Jansen *et al.* [4]. The study used generalized estimating equations (GEEs) to predict amyloid burden from age, diagnosis (control or MCI) and *APOE* $\epsilon 4$ status (defined as the presence or absence of any number of $\epsilon 4$ alleles). Age is centered at 70 years. In this study we only consider healthy controls and people with a diagnosis of MCI. The best model identified by Jansen *et al.* was:

$$Final\ model: A\beta \sim Age + Diagnosis + APOE + Age * Diagnosis + Age * APOE$$

We have used the estimates from this meta-analysis to inform the regression estimates of variables in this

study where possible. It can be shown that the estimates from GEEs are normally distributed [30] and hence we included these estimates as priors using a multivariate normal distribution (see Table 1). The PGRS had a non-informative multivariate normal prior with mean 0 and variance 100. We also created models where all variables had the non-informative Normal(0,100) priors for comparison.

Table 1: Informative prior distributions

Variable	**Estimate**	**SE**
Intercept	0.879	0.0967
Age	0.064	0.006
Diagnosis = Control	-0.964	0.0793
APOE status = 0	-1.493	0.0772
Age * *APOE* status = 0	-0.021	0.0079
Age * Diagnosis = Control	0.019	0.0081

SE = Standard error.

MCI is used as the reference diagnosis.

APOE status 1 (at least one $\epsilon 4$ allele) is used as the reference level.

177

The tuning parameter of each model was adjusted to achieve an acceptance rate in the Metropolis sampler of approximately 0.35. This tuning was performed over the 3,000 burn-in samples. This acceptance rate is slightly higher than advised in the literature as lower acceptance rates were causing reduced mixing [31].

2.4.3 Graphical User Interface

The analysis methods used in this study have been packaged into a user-friendly application through Rshiny [32]. The application is available to download from https://github.com/KHP-Informatics/bayesian-logistic-regression-r-shiny-app.git.

3 Results

3.1 PGRS

In the ADNI cohort the genetic data was imputed from 479,073 and 599,526 SNPs in ADNI 1 and ADNI 2 respectively, to 8,799,802 and 6,336,499 SNPs. In EDAR and DESCRIPA the data was increased from 619,609 SNPs to 5,409,779 by imputation. The PGRS was standardized by subtracting the mean and dividing by the standard deviation, per cohort.

3.2 Cohort demographics

The demographics given below are from the individuals in ADNI 1, ADNI 2, EDAR and DESCRIPA with CSF and GWAS data available.

In ADNI 1 (training data) there is 1 point difference in median MMSE between subjects with normal and abnormal Aβ (28 vs. 29). However, the larger sample size of ADNI 1 (compared to ADNI 2, EDAR and DESCRIPA) means this difference is statistically significant. Diagnosis and *APOE* genotype also show significant differences between groups, as we would expect. We see a significant difference (p-value < 0.05) in the PGRS between groups. Similar associations are seen in the data from EDAR and DESCRIPA with age also being nominally significantly associated with normal and abnormal Aβ. ADNI 2 data shows no significant difference in any demographic variable. This is likely to be due to the small sample size (N=43) of this population.

Table 2: Cohort demographics

	ADNI 1 (N=272)			**ADNI 2** (N=43)			**EDAR and DESCRIPA *** (N=127)		
Demographic	**Normal CSF Aβ** (N = 105)	**Abnormal CSF Aβ** (N = 167)	**P-value**	**Normal CSF Aβ** (N = 27)	**Abnormal CSF Aβ** (N = 16)	**P-value**	**Normal CSF Aβ** (N = 60)	**Abnormal CSF Aβ** (N = 67)	**P-value**
Median age [IQR]	74.1 [8.10]	75.1 [8.05]	0.837	67.8 [12.55]	73.75 [11.075]	0.087	66 [8.75]	69.55 [11.83]	0.047
Gender (%):			0.61			0.528			0.859
Female	36.2	39.5		51.9	37.5		46.7	49.3	
Male	63.8	60.5		48.1	62.5		53.3	50.7	
Median years in education [IQR]	16 [4]	16 [4]	0.955	16 [4]	16[4]	0.949	12 [7]	10 [5]	0.334
Median MMSE [IQR]	29 [3]	28 [3]	<0.001	29 [2]	27.5 [2.25]	0.1	28.5 [3]	27 [5]	0.024
Diagnosis (%)			<0.001			0.446			0.002
MCI	42.9	76.6		74.1	87.5		71.7	92.5	
CTL	57.1	23.4		25.9	12.5		28.3	7.5	
APOE status (%)			<0.001			0.343			0.004
0	84.8	38.3		66.7	50.0		66.7	40.3	
1	15.2	61.7		33.3	50.0		33.3	59.7	
Median PGRS [IQR]	-0.615 [2.196]	-0.560 [2.095]	0.018	-0.088 [1.362]	-0.221 [1.127]	0.763	-0.252 [1.387]	0.226 [1.266]	0.022

Kruskal Wallis chi-squared was used to test between normal and abnormal groups for continuous demographic variables.

Fishers exact was used to test between normal and abnormal groups for categorical demographic variables.

APOE status is 1 if an individuals genotype contains any $\epsilon 4$ alleles, and 0 otherwise.

IQR = Inter-quartile range; CSF = Cerebrospinal Fluid; MMSE = Mini Mental State Exam; MCI = Mild Cognitive Impairment; CTL = Control.

* One individual has missing MMSE and education information.

3.3 Data analysis

The test data indicated that, in most cases, the addition of informative priors on age, diagnosis and *APOE* genotype did not improve the predictive ability of models. Furthermore, no model including a PGRS showed higher accuracy than the basic models. The basic model with non-informative priors achieved the highest accuracies at 54% and 49% with a high sensitivity of 81% and 63%, but low specificity (23% and 41%). See Table 3 for full results.

Table 3: Test data results

Model	Informative priors?	Accuracy [95% CI]	Sensitivity	Specificity	AUC ROC
EDAR and DESCRIPA					
Demographics	No	0.535 [0.445; 0.624]	0.806	0.233	0.412
Demographics	Yes	0.543 [0.453; 0.632]	0.776	0.283	0.412
PGRS	No	0.457 [0.368; 0.547]	0.657	0.233	0.391
PGRS	Yes	0.394 [0.308; 0.484]	0.493	0.283	0.394
ADNI 2					
Demographics	No	0.488 [0.333; 0.645]	0.625	0.407	0.458
Demographics	Yes	0.488 [0.333; 0.645]	0.625	0.407	0.461
PGRS	No	0.465 [0.312; 0.623]	0.625	0.37	0.396
PGRS	Yes	0.442 [0.291; 0.601]	0.5	0.407	0.4

CI = Confidence Interval; AUC ROC = Area under the Receiver Operating Characteristic Curve;

PGRS = Polygenic Risk Score

209

4 Discussion

This study shows that the predictive ability of models including age, diagnosis and *APOE* genotype is not improved by the addition of a PGRS despite previous studies showing an association [13]. The PGRS used in this analysis was trained on a case/control endpoint. As GWA studies of amyloid endpoints become available the predictive ability of a PGRS trained using this information is likely to improve [33].

This paper presents the first analysis to use a Bayesian methodology in AD blood biomarker research. We aimed to inform estimates of well-researched risk variables (age, *APOE* status and control or MCI status) by including prior information from a large meta-analysis [4]. In this study we see that this approach does not improve predictive ability of models over those without informative prior information. However, we have only used one type of Bayesian methodology. There is a risk that if ill-fitting priors are used in combination with small sample sizes, model fit may be driven by the prior distributions. It is possible that the priors used here are not optimal as our population demographics are slightly different from those seen by Jansen *et al.* [4]. However, we believe informing model estimates with information from previous literature is likely to reduce false positives in biomarker studies of a small size. Bayesian analysis is one way of doing that; other methods should also be investigated.

The creation of the PGRS also created some limitations for this study. Firstly, the PGRS used here only utilizes common genetic variation. This is because rare genetic variants, such as *TREM2*, may not be significantly associated with disease in small populations. As larger studies become available inclusion of such variants should be investigated. Further, the PGRS was created using a simple additive method. This simplistic method is likely to be sub-optimal and as new methodologies for creating PGRS become available they should be investigated in this setting. Additionally, the platforms used to measure genetics differed between the discovery and test cohorts used in this study. It is well documented that there can be inconsistencies between *omics* platforms which could have contributed to the reduced predictive ability seen here [34, 35, 36, 37].

It is important to bear in mind in all further research that differences in normal/abnormal CSF Aβ cut-offs can make replication and research in other cohorts difficult. Work is being done to investigate these values with an aim of standardization [38]. This study also has the limitation that the IGAP data (used to generate the PGRS) is not independent of the training and test data, as ADNI was included in IGAP. However, we believe the benefit of larger sample size outweighs this. Finally, the cohorts used in this study are still of a relatively small size. Although we tried to address this through the use of Bayesian methods,

studies of larger sample sizes will be vital for further investigations of blood based biomarkers of Aβ.

This study also presents opportunities for further work. Firstly, it is possible that markers identified in case/control studies are associated with other AD related phenotypes such as tau burden. If large meta-analyses of the risk factors affecting tau burden, and other endophenotypes, become available it would be interesting to perform similar Bayesian analysis on these alternative endpoints. Secondly, in the setting of preventative clinical trials the most common way to measure brain amyloid burden is through the use of a PET scan. In this study we have used CSF. However, it has been shown that measurements from CSF and PET are highly correlated meaning markers identified for one can reasonably be tested to see whether they are transferable to the other. In this study the use of CSF allowed us to maximize the sample size with measurements of amyloid and genetics. If promising blood markers of CSF measures are identified, future studies should perform similar analysis using a measurement of Aβ derived from a PET scan. Furthermore, the effect of a measure of brain reserve on model accuracy could be investigated. It is well-known that some people with high levels of brain amyloid burden at autopsy show no cognitive deficit during their lifetime. It has been shown in recent studies that levels of 'brain reserve' may be driving the difference between people who have high levels of pathology and no symptoms and those who show symptoms [39]. This is motivating a theory that increased brain reserve may prevent development of symptoms even if pathology is present. Although brain reserve itself may be hard to quantify it may be possible to measure and model associated lifestyle, environmental and psychological factors such as social networks.

Although there are further candidates and alternative methods to be considered in the search for a blood biomarker of amyloid burden it is imperative that appropriate populations are used. As in this work, the search in asymptomatic individuals is likely to have the biggest impact for enrichment of clinical trials. Furthermore, rigorous testing of biomarkers is essential. Without replication it is probable that model performance is overestimated.

5 Conclusions

This paper presents the first analysis to use a Bayesian methodology in the search for a blood biomarker of AD. We see that the Bayesian approach does not improve predictive ability in this setting, and that *omics* measurements do not improve the predictive ability of models above that of demographics alone. We have been unable to demonstrate any additional benefit over age diagnosis and *APOE* genotype by including a case/control PGRS when predicting amyloid positivity in subjects without a clinical diagnosis of AD or

other dementia.

6 Acknowledgements

This work represents independent research part funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. This study was supported by researchers at the National Institute for Health Research University College London Hospitals Biomedical Research Centre. The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement number 115372, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution. Steven Kiddle is supported by an MRC Career Development Award in Biostatistics (MR/L011859/1). Nicola Voyle is funded by the Alzheimer's Society. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

The authors would like to thank Dr. Steven Hill and Dr. Robert Goudie from the MRC Biostatistics Unit for their guidance in Bayesian methods.

This work was also supported by awards to establish the Farr Institute of Health Informatics Research at UCL partners, London, from the Medical Research Council, Arthritis Research UK, British Heart Foundation, Cancer Research UK, Chief Scientist Office, Economic and Social Research Council, Engineering and Physical Sciences Research Council, National Institute for Health Research, National Institute for Social Care and Health Research, and Wellcome Trust (grant MR/K006584/1).

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimers Association; Alzheimers Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research and Development, LLC.; Johnson and Johnson Pharmaceutical Research and Development LLC.; Lumosity; Lundbeck; Merck and Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals

Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

References

[1] Querfurth H and LaFerla F. Mechanisms of Disease: Alzheimer's Disease. New England Journal of Medicine 2011;362: 329–344.

[2] Villemagne VL, Pike KE, Chételat G, Ellis KA, Mulligan RS, Bourgeat P *et al.* Longitudinal assessment of Aβ and cognition in aging and Alzheimer disease. Annals of neurology 2011;69(1): 181–192.

[3] Jack CR, Knopman DS, Jagust WJ, Petersen RC, Weiner MW, Aisen PS *et al.* Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. The Lancet Neurology 2013;12(2): 207–216.

[4] Jansen WJ, Ossenkoppele R, Knol DL, Tijms BM, Scheltens P, Verhey FRJ *et al.* Prevalence of Cerebral Amyloid Pathology in Persons Without Dementia. JAMA 2015;313(19): 1915–1924.

[5] Gatz M, Pedersen NL, Berg S, Johansson B, Johansson K, Mortimer Ja *et al.* Heritability for Alzheimer's disease: the study of dementia in Swedish twins. The journals of gerontology Series A, Biological sciences and medical sciences 1997;52A(2): M117–M125.

[6] Pedersen NL, Gatz M, Berg S and Johansson B. How Heritable Is Alzheimer's Disease Late in Life? Findings from Swedish Twins. Annals of Neurology 2004;55: 180–185.

[7] Ridge PG, Hoyt KB, Boehme K, Mukherjee S, Crane PK, Haines JL *et al.* Assessment of the genetic variance of late-onset Alzheimer's disease. Neurobiology of aging 2016;41: 200.e13–200.e20. doi: 10.1016/j.neurobiolaging.2016.02.024.

[8] Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nature Genetics 2013; 45(12): 1452–1458.

[9] Lee SH, Harold D, Nyholt DR, Consortium A, Consortium IE, disease Consortium G *et al.* Estimation and partitioning of polygenic variation captured by common SNPs for Alzheimer's disease, multiple sclerosis and endometriosis. Human Molecular Genetics 2013;22(4): 832–841.

[10] Ridge PG, Mukherjee S, Crane PK, Kauwe JSK and Consortium ADG. Alzheimer's Disease: Analyzing the Missing Heritability. PLOS ONE 2013;8(11): e79771.

[11] Ruiz A, Heilmann S, Becker T, Hernández I, Wagner H, Thelen M *et al.* Follow-up of loci from the International Genomics of Alzheimer' s Disease Project identifies TRIP4 as a novel susceptibility gene. Translational Psychiatry 2014;4(January): e358. doi:10.1038/tp.2014.2.

[12] Cruchaga C, Ebbert MTW and Kauwe JSK. Genetic Discoveries in AD Using CSF Amyloid and Tau. Current Genetic Medicine Reports 2014;2(1): 23–29.

[13] Sabuncu MR, Buckner RL, Smoller JW, Lee PH, Fischl B, Sperling RA *et al.* The Association between a Polygenic Alzheimer Score and Cortical Thickness in Clinically Normal Subjects. Cerebral Cortex 2012;22(11): 2653–2661.

[14] Sleegers K, Bettens K, De Roeck A, Van Cauwenberghe C, Cuyvers E, Verheijen J *et al.* A 22-single nucleotide polymorphism Alzheimer risk score correlates with family history, onset age, and cerebrospinal fluid Aβ42. Alzheimer's & Dementia 2015;11(12): 1452–1460.

[15] Escott-Price V, Sims R, Bannister C, Harold D, Vronskaya M, Majounie E *et al.* Common polygenic variation enhances risk prediction for Alzheimer's disease. Brain 2015;138(12): 3673–3684.

[16] Ossenkoppele R, Jansen WJ, Rabinovici GD, Knol DL, van der Flier WM, van Berckel BNM *et al.* Prevalence of Amyloid PET Positivity in Dementia Syndromes. JAMA 2015;313(19): 1911–1939.

[17] Barnett JH, Blackwell A, Scheltens P, Waldemar G, Johannsen P, Tsolaki M *et al.* Cognitive function and cognitive change in dementia, mild cognitive impairment, and healthy aging: The EDAR study. Alzheimer's & Dementia 2010;6(4): S127.

[18] Visser PJ, Verhey FRJ, Boada M, Bullock R, De Deyn PP, Frisoni GB *et al.* Development of Screening Guidelines and Clinical Criteria for Predementia Alzheimer's Disease. Neuroepidemiology 2008;30(4): 254–265.

[19] Visser PJ, Verhey F, Knol DL, Scheltens P, Wahlund LO, Freund-Levi Y *et al.* Prevalence and prognostic value of CSF markers of Alzheimer's disease pathology in patients with subjective cognitive

impairment or mild cognitive impairment in the DESCRIPA study: a prospective cohort study. The Lancet Neurology 2009;8(7): 619–627.

[20] Goldstein JI, Crenshaw A, Carey J, Grant GB, Maguire J, Fromer M *et al.* Zcall: A rare variant caller for array-based genotyping. Bioinformatics 2012;28(19): 2543–2545.

[21] Saykin AJ, Shen L, Foroud TM, Potkin SG, Swaminathan S, Kim S *et al.* Alzheimer's Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: Genetics core aims, progress, and plans. Alzheimer's & Dementia 2010;6(3): 265–273.

[22] Coleman JRI, Euesden J, Patel H, Folarin AA, Newhouse S and Breen G. Quality control , imputation and analysis of genome-wide genotyping data from the Illumina HumanCoreExome microarray. Briefings in Functional Genomics 2015;1–7.

[23] Auton A, Abecasis GR, Altshuler DM, Durbin RM, Bentley DR, Chakravarti A *et al.* A global reference for human genetic variation. Nature 2015;526(7571): 68–74.

[24] R Core Team. R: A Language and Environment for Statistical Computing, 2014.

[25] Kuhn M, Wing J, Weston S, Williams A, Keefer C, Engelhardt A *et al.* caret: Classification and Regression Training, 2015.

[26] Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC *et al.* pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011;12(1): 77.

[27] Sing T, Sander O, Beerenwinkel N and Lengauer T. ROCR: visualizing classifier performance in R. Bioinformatics 2005;21(20): 3940–3941.

[28] Euesden J, Lewis CM and O'Reilly PF. PRSice: Polygenic Risk Score software. Bioinformatics 2015; 31(9): 1466–1468.

[29] Martin A and Quinn K. Applied Bayesian inference in R using MCMCpack. R News 2006;6(March): 2–7.

[30] Ziegler A, Kastner C and Blettner M. The generalised estimating equations: an annotated bibliography. Biometrical Journal 1998;40: 115–140.

[31] Roberts GO, Gelman A and Gilks WR. Weak convergence and optimal scaling of random walk Metropolis algorithms. Annals of Applied Probability 1997;7(1): 110–120.

[32] Chang W, Cheng J, Allaire J, Xie Y and McPherson J. shiny: Web Application Framework for R, 2015.

[33] Dudbridge F. Power and Predictive Accuracy of Polygenic Risk Scores. PLoS Genetics 2013;9(3): e1003348.

[34] Dobbin KK, Beer DG, Meyerson M, Yeatman TJ, Gerald WL, Jacobson JW *et al.* Interlaboratory comparability study of cancer gene expression analysis using oligonucleotide microarrays. Clin Cancer Res 2005;11(2 Pt 1): 565–572.

[35] Leek JT, Scharpf RB, Bravo HC, Simcha D, Langmead B, Johnson WE *et al.* Tackling the widespread and critical impact of batch effects in high-throughput data. Nature Reviews Genetics 2010;11(10): 733–739.

[36] McShane LM, Cavenagh MM, Lively TG, Eberhard DA, Bigbee WL, Williams PM *et al.* Criteria for the use of omics-based predictors in clinical trials. Nature 2013;502(7471): 317–320.

[37] O'Bryant SE, Lista S, Rissman RA, Edwards M, Zhang F, Hall J *et al.* Comparing biological markers of Alzheimer's disease across blood fraction and platforms: Comparing apples to oranges. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring 2015;3: 27–34.

[38] Zwan MD, Rinne JO, Hasselbalch SG, Nordberg A, Lleo A, Herukka SK *et al.* Use of amyloid-PET to determine cutpoints for CSF markers: A multicenter study. American Academy of Neurology 2015; 86(1): 50–58.

[39] Bauckneht M, Picco A, Nobili F and Morbelli S. Amyloid positron emission tomography and cognitive reserve. World Journal of Radiology 2015;7(12): 475–483.

Chapter 6

Genetic risk as a marker of amyloid-β and tau burden in cerebrospinal fluid

Journal of Alzheimer's Disease 1 (2016) 1–11
DOI 10.3233/JAD-160707
IOS Press

Genetic Risk as a Marker of Amyloid-β and Tau Burden in Cerebrospinal Fluid

Nicola Voyle[a], Hamel Patel[a], Amos Folarin[a], Stephen Newhouse[a], Caroline Johnston[a], Pieter Jelle Visser[b,c], Richard J.B. Dobson[a,d,e,1,*] and Steven J. Kiddle[a,f,1,*] the EDAR and DESCRIPA study groups and the Alzheimer's Disease Neuroimaging Initiative[2]
[a]*MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK*
[b]*Department of Neurology and Alzheimer Center, VU University Medical Center, Amsterdam, The Netherlands*
[c]*Helen Wills Neuroscience Institute, University of California, Berkeley, CA, USA*
[d]*NIHR Biomedical Research Centre for Mental Health and Biomedical Research Unit for Dementia at South London and Maudsley NHS Foundation, London, UK*
[e]*Farr Institute of Health Informatics Research, UCL Institute of Health Informatics, University College London, London, UK*
[f]*MRC Biostatistics Unit, Cambridge Biomedical Campus, Cambridge Institute of Public Health, Forvie Site, Robinson Way, Cambridge, UK*

Handling Associate Editor: Gary Arendash

Accepted 28 September 2016

Abstract.
Background: The search for a biomarker of Alzheimer's disease (AD) pathology (amyloid-β (Aβ) and tau) is ongoing, with the best markers currently being measurements of Aβ and tau in cerebrospinal fluid (CSF) and via positron emission tomography (PET) scanning. These methods are relatively invasive, costly, and often have high screening failure rates. Consequently, research is aiming to elucidate blood biomarkers of Aβ and tau.
Objective: This study aims to investigate a case/control polygenic risk score (PGRS) as a marker of tau and investigate blood markers of a combined Aβ and tau outcome for the first time. A sub-study also considers plasma tau as markers of Aβ and tau pathology in CSF.
Methods: We used data from the EDAR*, DESCRIPA**, and Alzheimer's Disease Neuroimaging Initiative (ADNI) cohorts in a logistic regression analysis to investigate blood markers of Aβ and tau in CSF. In particular, we investigated the extent to which a case/control PGRS is predictive of CSF tau, CSF amyloid, and a combined amyloid and tau outcome. The predictive ability of models was compared to that of age, gender, and APOE genotype ('basic model').
Results: In EDAR and DESCRIPA test data, inclusion of a case/control PGRS was no more predictive of Aβ, and a combined Aβ and tau endpoint than the basic models (accuracies of 66.0%, and 73.3% respectively). The tau model saw a small increase in accuracy compared to basic models (59.6%). ADNI 2 test data also showed a slight increase in accuracy for the Abeta model when compared to the basic models (61.4%).
Conclusion: We see some evidence that a case/control PGRS is marginally more predictive of Aβ and tau pathology than the basic models. The search for predictive factors of Aβ and tau pathologies, above and beyond demographic information,

[1]These authors are joint senior authors.

[2]Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (http://adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data, but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at http://adni.loni.usc.edu/wp–content/uploads/how to apply/ADNI Acknowledgement List. pdf.

*Correspondence to: Richard J.B. Dobson and Steven J. Kiddle, SGDP Centre, IoPPN, 16 De Crespigny Park, Denmark Hill, London SE5 8AF, UK. Tel.: +44 020 7848 0924; Fax: +44 020 7848 0866; E-mails: richard.j.dobson@kcl.ac.uk (Richard J.B. Dobson), steven.kiddle@kcl.ac.uk (Steven J. Kiddle).

ISSN 1387-2877/16/$35.00

is still ongoing. Better understanding of AD risk alleles, development of more sensitive assays, and studies of larger sample size are three avenues that may provide such factors. However, the clinical utility of possible predictors of brain Aβ and tau pathologies must also be investigated.

*'Beta amyloid oligomers in the early diagnosis of AD and as marker for treatment response'

**'Development of screening guidelines and criteria for pre-dementia Alzheimer's disease'

Keywords: Alzheimer's disease, biomarker, blood, multi-modal, polygenic risk score, tau

INTRODUCTION

The hallmark pathologies of Alzheimer's disease (AD) are amyloid-β (Aβ) plaques and phosphorylated tau tangles in the brain. Although diagnostic criteria for AD focus on pathological evidence of Aβ, tau levels in cerebrospinal fluid (CSF) are also considered [1]. Jack et al. provide a theoretical model for the progression of AD based on existing evidence that is consistent with the popular amyloid cascade hypothesis [2, 3]. This hypothesis states that the build up of tau is triggered by increasing levels of Aβ in the brain.

The search for a biomarker of AD pathologies (tau and Aβ) is ongoing with the best markers currently being measurements of tau and Aβ in the CSF and via positron emission tomography (PET) scanning. These methods are relatively invasive and often have high screening failure rates meaning a high proportion of individuals that are scanned have low levels of these pathologies. Additionally, PET scanning in particular is an expensive procedure costing around $3,000 per scan [4]. There are only approximately 2,380 (http://www.imvinfo.com, August 2016) PET scanners in the United States meaning access to facilities is limited [5]. The lumbar puncture needed to access CSF is also considered a high-risk procedure in many western countries [6]. To address these shortcomings, research is aiming to elucidate blood biomarkers of AD pathologies (Aβ and tau) [7]. One motivation for discovering a blood-based biomarker of AD pathology comes from clinical trial recruitment. For example, when recruiting into a trial of an anti-amyloid therapeutic, a blood-based biomarker of Aβ could act as a filtering step to identify individuals with abnormal levels of Aβ before performing a confirmatory PET scan or lumbar puncture. As such a test is likely to be cost-effective, we could reduce the cost of screening while also reducing the number of individuals subject to the invasive lumbar puncture and PET scanning procedures. So far, research into a blood marker has largely focused on Aβ brain burden, rather than tau pathology. However, some studies have investigated genetic markers of tau [8]. This study aims to further this research by investigating markers of a combined Aβ and tau outcome for the first time.

Genome wide association studies (GWAS) of AD to date have identified over 20 risk loci explaining approximately 16–33% of genetic variability in the disease [9–13]. Compared to the predicted heritability of 50–70%, this is fairly low [13–15]. Modern technologies including next generation sequencing and the development of high coverage, exome chips are beginning to address this issue of missing heritability. Meanwhile polygenic risk scores (PGRS) are aiming to combine genetic risk from variants of lower effect size [16]. To date, PGRS in AD have only been trained on case/control endpoints as GWAS studies of pathological outcomes are relatively small (for example, N = 1,269 [17]). Studies have shown case/control PGRS are associated with AD-related phenotypes but few have investigated their predictive ability in test data [18, 19]. When we studied the predictive ability of a PGRS for Aβ burden, we saw that the case/control PGRS used was no more predictive than demographics (age, gender) and APOE genotype alone (Voyle et al., in submission). However, we hypothesize in the current study that the PGRS may be more predictive of a combined Aβ and tau outcome; a more 'AD-like' phenotype.

This study aims to investigate the predictive ability of a case/control PGRS on CSF Aβ and tau. A substudy will also consider plasma tau as a predictor. We use individuals from the Alzheimer's Disease Neuroimaging Initiative (ADNI), EDAR, and DESCRIPA studies to investigate these associations and compare all models to those of demographics (age, gender) and APOE genotype [20, 21]. We hypothesize that blood markers of AD will perform better when predicting a combined pathology (tau and Aβ) endpoint over tau and Aβ individually. The combined endpoint should be more representative of an 'AD-like' phenotype.

MATERIALS AND METHODS

Cohorts

EDAR was a prospective, longitudinal study which aimed to examine and evaluate biomarkers of early AD and treatment response [21]. In particular, the study focused on Aβ oligomers and the effect of genetic variants on these oligomers. For more information see http://www.edarstudy.eu. Our access to samples and clinical and phenotypic information from the EDAR study was enabled by the European Medical Information Framework and has been previously described (Voyle et al. in submission).

DESCRIPA was also a prospective, multi-center study. It was coordinated by the European AD Consortium and focused on collecting data from non-demented subjects with the aim of developing screening guidelines and clinical criteria for AD. Further details of this study can be found in Visser et al. [20].

ADNI is a longitudinal cohort study aiming to validate the use of biomarkers in AD clinical trials and diagnosis. Data used in the preparation of this article were obtained from the ADNI database (http://adni.loni.usc.edu). The ADNI study was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether biological markers and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and AD. ADNI was approved by the institutional review boards of all participating institutions, and written informed consent was obtained for all participants. This study uses data from ADNI 1 and the ADNI 2 and ADNI GO sub-groups, referred to as ADNI 2 from here onwards.

Genetics

Samples from EDAR and DESCRIPA were genotyped on the Illumina HumanOmniExpressExome-8v1.2 BeadChip (N = 336) [22]. The data has previously been reported on by Voyle et al. (in submission). The HumanOmniExpressExome BeadChip has been optimized to tag SNPs that capture a large amount of common genetic variation. In total, the chip contains 960,919 markers of which over 273,000 correspond to functional exomic markers. ADNI 1 samples were run on the Human610-Quad BeadChip (N = 818), which has since been discontinued. The chip provides coverage of 924,000 randomly selected SNPs. ADNI 2 and ADNI GO samples were run on the Illumina HumanOmniExpress BeadChip (N = 432). This chip is similar to the HumanOmniExpressExome BeadChip used in the EDAR and DESCRIPA studies but does not include the exomic markers. In total, the chip contains 713,599 markers. Details of the genotyping protocols followed in ADNI are given elsewhere [23]. Details of the data processing are briefly outlined below.

The cohorts were subject to quality control and imputation separately, as described in Coleman et al. [24]. In short, the data was filtered to ensure a minor allele frequency of greater than 5% for all SNPs before removal of rare variants and subjects with high levels of missing data. SNPs that differed significantly ($p < 0.00001$) from the Hardy-Weinberg equilibrium were removed. The data was pruned for SNPs in linkage disequilibrium and for genetically similar individuals. Finally, the data was imputed using reference files from the 1000 Genomes Project [25].

CSF Measurements

This study focuses on total tau (tTau) and Aβ measurements in CSF. The analysis considered three endpoints: dichotomized Aβ, dichotomized tTau, and a binary representation of overall pathology. For the latter analysis, referred to as total CSF burden, each cohort was reduced to those individuals with normal Aβ and normal tTau, or abnormal Aβ and abnormal tTau. The distributions of tTau and Aβ were similar between cohorts in terms of shape but not in absolute value. Consequently, this work focused solely on a dichotomized outcome.

EDAR

Details of CSF collection and analysis can be found at http://www.edarstudy.eu. In brief, CSF measurements were collected using the Alzbio3 Luminex assay in one batch at the end of the study. CSF Aβ and tTau measurements were dichotomized at the previously published thresholds of 389 pg/ml and 98 pg/ml, respectively.

DESCRIPA

Details of CSF measurements in DESCRIPA have been described elsewhere [22]. In brief, CSF measurements were analyzed in one laboratory and collected using single-parameter ELISA kits

(Innogenetics, Ghent, Belgium). CSF Aβ and tTau samples were dichotomized using the previously published thresholds of 550 pg/ml and 375 pg/ml, respectively.

ADNI

For ADNI, datasets used to extract CSF measures of Aβ and tau were chosen to maximize sample size. The dataset 'UPENNBIOMK2' was used for ADNI 1 and 'UPENNBIOMK6' for ADNI 2 and ADNI GO. Both datasets contain CSF measurements collected using the xMAP Luminex platform and Innogenetics immunoassay kits. The CSF measures of amyloid and tTau were dichotomized at the previously published thresholds (192 pg/ml and 93 pg/ml, respectively).

Plasma Tau (ADNI 1 only)

Plasma tau was investigated as a potential blood biomarker of Aβ and tau in a sub-study. ADNI 1 was the only cohort with such data available. Plasma tau was analyzed by the Single Molecule Array (SIMOA) technique and the Human total tau assay using a combination of monoclonal antibodies. Samples with a plasma tau concentration below the lower limit of quantification (<1.0 pg/ml) were removed (N = 36). Outliers, identified as values greater than 6 standard deviations from the mean, were removed (N = 2) and the data was subject to a natural logarithm transformation.

Statistical analysis

All statistical analysis was performed in R version 3.1.1 [26].

The three endpoints of interest (dichotomized amyloid, dichotomized tTau, and total CSF burden) were modeled using logistic regression models covarying for age, gender, and APOE genotype. An individuals APOE genotype was coded as 1 if at least one $\varepsilon 4$ allele was present and 0 otherwise. Models were also built using only the demographic variables age, gender, and APOE genotype, for comparison. Throughout this study these models are referred to as 'basic models'.

Two analyses were performed to study these three endpoints of interest:

1. PGRS: Models were built in ADNI 1 data (Aβ and tau N = 363; Total CSF burden N = 244) and tested in data from EDAR and DESCRIPA (Aβ and tau N = 250; Total CSF burden N = 135) and ADNI 2 (Aβ and tau N = 44; Total CSF burden N = 37).
2. PGRS and plasma tau: Models were built and tested in ADNI 1 using 5 fold cross-validation (Aβ and tau N = 323; Total CSF burden N = 219).

Table 1
Sample sizes

PGRS analysis			
Outcome	ADNI 1 (N)	EDAR and DESCRIPA (N)	ADNI 2 (N)
Dichotomized amyloid or tTau	363	250	44
Total CSF burden	244	135	37
PGRS and plasma tau analysis			
Outcome	ADNI 1 (N)		
Dichotomized amyloid or tTau	323		
Total CSF burden	323		

The sample sizes of each dataset used in these analyses are given in Table 1.

All models including a PGRS co-varied for the first three genetic principal components to account for population structure. The predictive ability of each model was quantified using accuracy, sensitivity, specificity and area under the receiver operating characteristic (ROC) curve [27–29].

PGRS

PGRS were created within each cohort using PRSice [30]. Effect sizes from stage 1 of the International Genomics of Alzheimer's Project (IGAP) case/control GWAS were used as the weights to generate the risk score [9]. We used 0.5 as the *p*-value threshold for inclusion in the PGRS. This threshold showed the most significant association with case/control diagnosis in the large IGAP PGRS study [31]. The genetic region coding for APOE was removed from all scores and included as a separate covariate due to its strong influence. The PGRS was standardized by subtracting the mean and dividing by the standard deviation, per cohort. This aims to account for the scores including different SNPs due to availability on the different SNP chips. It is important to note that APOE genotype is included as a covariate in modeling and not in the PGRS. Therefore, the PGRS is only exploring variation above and beyond APOE. This is different to the focus of the IGAP PGRS study which included APOE within the PGRS [31]. We have chosen this approach to investigate

Table 2
Cohort demographics – Dichotomized Aβ

Demographic	ADNI 1 (N = 363)			ADNI 2 (N = 44)			EDAR and DESCRIPA* (N = 250)		
	Normal CSF Aβ (N = 112)	Abnormal CSF Aβ (N = 251)	*p*-value	Normal CSF Aβ (N = 27)	Abnormal CSF Aβ (N = 17)	*p*-value	Normal CSF Aβ (N = 99)	Abnormal CSF A β (N = 151)	*p*-value
Median age [IQR]	74.4 [8.35]	75.3 [8.55]	0.825	67.8 [12.55]	75.2 [11.1]	0.049	66 [12]	69 [13]	0.136
Gender (%)			0.487			0.359			0.052
Female	36.6	40.6		51.9	35.3		38.4	51.7	
Male	63.4	59.4		48.1	64.7		61.6	48.3	
Median years in education [IQR]	16 [4.25]	16 [4]	0.748	16 [4]	16 [4]	0.825	11.5 [8]	10 [7]	0.133
Median MMSE [IQR]	29 [3]	26 [4]	<0.001	29 [2]	27 [3]	0.058	27.5 [4]	26 [5]	0.009
Diagnosis (%)			<0.001			0.247			0.001
Dementia	6.3	33.5		0	5.9		34.3	45.0	
MCI	40.1	51.0		74.1	82.4		43.4	41.1	
SCI	0	0		0	0		5.1	10.6	
CTL	53.6	15.5		25.9	11.7		17.2	3.3	
APOE status (%)			<0.001			0.225			<0.001
0	85.7	33.5		66.7	47.1		74.7	38.4	
1	14.3	66.5		33.3	52.9		25.3	61.6	

Kruskal Wallis chi-squared was used to test between normal and abnormal groups for continuous demographic variables. Fishers exact was used to test between normal and abnormal groups for categorical demographic variables. APOE status is 1 if an individual's genotype contains any ε4 alleles, and 0 otherwise. IQR, Inter-quartile range; CSF, cerebrospinal fluid; MMSE, Mini-Mental State Exam; MCI, mild cognitive impairment; SCI, subjective cognitive impairment; CTL, control. *One individual has missing diagnosis, 2 have missing education information, and 3 have missing MMSE.

whether information from a SNP chip adds anything above APOE. Furthermore, generally, APOE is not well measured on SNP chips so APOE genotype is determined by targeted genotyping.

The datasets used for model testing (EDAR and DESCRIPA) were not included in the IGAP study. Furthermore, although some samples from ADNI were included in IGAP, the majority were only included in the training, not testing, stages.

RESULTS

Cohort demographics

Table 2 shows demographics against normal and abnormal CSF Aβ while Table 3 is against CSF tTau. Demographics for the sub-population used in the total CSF burden analysis are given in Supplementary Table 1. Demographics for the sub-group of ADNI 1 individuals with plasma tau measurements is given in Supplementary Table 2.

In ADNI 1 training data and EDAR/DESCRIPA test data, there is a significant difference between normal and abnormal Aβ in MMSE, diagnosis, and APOE genotype as we would expect. Similar associations are also seen with dichotomized tTau. The smaller sample size of the ADNI 2 test data (N = 44) is likely to have driven the lack of association in this group of individuals.

Data analysis

Genetic risk

The PGRS was not significant in any of the logistic regression models (*p*-values of 0.995, 0.929, and 0.796 for tTau, Aβ, and total CSF burden respectively). The inclusion of the PGRS marginally improved the predictive ability of tTau models over the basic models. The accuracy of the Abeta model was also marginally improved by inclusion of the PGRS in ADNI 2 test data. The models of total CSF burden had the highest accuracies (72% and 65%). See Table 4 for full results.

Genetic risk and plasma tau

When modeling CSF tTau no model outperformed the basic model at an accuracy of 66%. Similarly, when modeling total CSF burden the inclusion of a PGRS and plasma tau did not improve predictive ability above the basic model (77%). The only model to see an improvement over the basic model was when modeling Aβ. Inclusion of plasma tau measurements marginally improved accuracy from 71% to 74% and the area under the ROC curve from 0.658 to 0.697. See Table 5 for full results.

DISCUSSION

The aim of this study was to investigate blood biomarkers that may be predictive of AD pathologies

Table 3
Cohort demographic - Dichotomized total tau

Demographic	ADNI 1 (N = 363)			ADNI 2 (N = 44)			EDAR and DESCRIPA* (N = 250)		
	Normal CSF tau (N = 203)	Abnormal CSF tau (N = 160)	*p*-value	Normal CSF tau (N = 30)	Abnormal CSF tau (N = 14)	*p*-value	Normal CSF tau (N = 150)	Abnormal CSF tau (N = 100)	*p*-value
Median age [IQR]	75.2 [8]	74.45 [10]	0.642	67.85 [10.7]	76.75 [12.68]	0.051	66.55 [13]	70.5 [12]	0.032
Gender (%)			0.234			>0.999			0.094
Female	36.5	43.1		46.7	42.9		42.0	53.0	
Male	63.5	56.9		53.3	57.1		58.0	47.0	
Median years in education [IQR]	16 [4]	16 [5]	0.064	16 [4]	16 [4]	0.7	10 [6]	9 [8]	0.032
Median MMSE [IQR]	28 [3]	26 [4]	<0.001	29 [1.75]	27 [3.75]	0.153	28 [4]	26 [4]	<0.001
Diagnosis (%)			<0.001			0.38			<0.001
Dementia	15.8	36.9		0	7.1		26.7	62.0	
MCI	45.8	50.0		76.7	78.6		47.3	34.0	
SCI	0	0		0	0		12.7	2.0	
CTL	38.4	13.1		23.3	14.3		13.3	2.0	
APOE status (%)			<0.001			0.191			0.001
0	63.5	31.9		66.7	42.9		59.3	43.0	
1	36.5	43.1		33.3	57.1		40.7	57.0	

Kruskal Wallis chi-squared was used to test between normal and abnormal groups for continuous demographic variables. Fishers exact was used to test between normal and abnormal groups for categorical demographic variables. APOE status is 1 if an individual's genotype contains any ε4 alleles, and 0 otherwise. IQR, inter-quartile range; CSF, cerebrospinal fluid; MMSE, Mini-Mental State Exam; MCI, mild cognitive impairment; SCI, subjective cognitive impairment; CTL, control. *One individual has missing diagnosis, 2 have missing education information, and 3 have missing MMSE.

Table 4
PGRS: Test data results

Outcome	Model	Accuracy	Sensitivity	Specificity	AUC ROC
EDAR and DESCRIPA					
tTau	Demographics only	0.584	0.570	0.593	0.625
Aβ	Demographics only	0.660	0.788	0.465	0.686
Total CSF burden	Demographics only	0.733	0.750	0.716	0.801
tTau	Demographics and PGRS	0.596	0.560	0.620	0.609
Aβ	Demographics and PGRS	0.616	0.768	0.384	0.692
Total CSF burden	Demographics and PGRS	0.719	0.702	0.702	0.748
ADNI 2					
tTau	Demographics only	0.636	0.571	0.667	0.655
Aβ	Demographics only	0.546	0.765	0.407	0.527
Total CSF burden	Demographics only	0.649	0.583	0.680	0.693
tTau	Demographics and PGRS	0.659	0.571	0.700	0.671
Aβ	Demographics and PGRS	0.614	0.647	0.593	0.590
Total CSF burden	Demographics and PGRS	0.649	0.583	0.680	0.683

AUC ROC, area under the receiver operating curve; PGRS, polygenic risk score.

(Aβ and tau). We modeled levels of Aβ and tau in CSF using a PGRS and a sub-study considered measurements of tau in blood plasma. We also studied a total CSF burden endpoint for individuals with abnormal tau and Aβ, or normal tau and Aβ. The results shown here highlight that a case/control PGRS and plasma tau do not substantially outperform demographics (age, gender) and APOE genotype. The highest model accuracies were seen when modeling the total CSF burden phenotype.

Several studies have focused on identification of blood biomarkers of Aβ [17, 32, 33]. However, few have achieved successful replication. The hypothesis tested in this study was that a combined tau and Aβ endpoint would be closer to an AD case/control phenotype and consequently easier to predict. This study supports this hypothesis, although the improvements in accuracy are minimal. Further, this accuracy of prediction was achieved by the basic model and the two blood biomarkers investigated, a case/control PGRS

Table 5
PGRS and plasma tau: Five fold cross-validation results

Outcome	Model	Accuracy	Sensitivity	Specificity	AUC ROC
tTau	Demographics only	0.656	0.691	0.628	0.659
tTau	Demographics and PGRS	0.638	0.665	0.617	0.641
tTau	Demographics and plasma tau	0.650	0.587	0.700	0.644
tTau	Demographics, PGRS and plasma tau	0.653	0.608	0.689	0.649
Aβ	Demographics only	0.709	0.787	0.530	0.658
Aβ	Demographics and PGRS	0.697	0.769	0.532	0.650
Aβ	Demographics and plasma tau	0.743	0.813	0.582	0.697
Aβ	Demographics, PGRS and plasma tau	0.725	0.813	0.523	0.668
Total CSF burden	Demographics only	0.772	0.742	0.816	0.779
Total CSF burden	Demographics and PGRS	0.763	0.735	0.804	0.769
Total CSF burden	Demographics and plasma tau	0.758	0.765	0.748	0.756
Total CSF burden	Demographics, PGRS and plasma tau	0.772	0.780	0.756	0.768

AUC ROC, area under the receiver operating curve; PGRS, polygenic risk score.

and plasma tau, did not improve model accuracy in the majority of cases.

Firstly, this could be explained by the PGRS being trained on a case/control endpoint. As individuals can often be misdiagnosed with AD the case/control phenotype can be misleading. GWAS are beginning to be large enough to detect risk SNPs associated with Aβ and tau. When sample sizes in these studies increase further they should be used to train PGRS. Intuitively, they may achieve improved predictive ability than models based on a case/control PGRS. Furthermore, due to the relatively small sample size of this study, the PGRS only utilizes common variation excluding loci such as TREM2. As larger studies become available the inclusion of such rarer variants in genetic risk scoring should be considered.

In the sub-study, it is interesting that plasma tau is no more predictive of CSF tau than age, gender, and APOE genotype. Furthermore, plasma tau achieves a higher accuracy when modeling CSF Aβ. The lack of ability to predict CSF tau indicates the need for further assay development to detect still more sensitive measurements. Furthermore, research into the permeability of the blood-brain barrier will help theorize as to how much pathology signal from the brain and CSF is likely to be seen in blood.

This study has shown the importance of replication in independent datasets; models that perform well in training data often do not replicate. It is particularly important to test replicability when standardized protocols for assays do not exist [34]. For example, this analysis highlights the difference between EDAR, DESCRIPA, and ADNI in the assays and cut-offs used to define high and low pathology burden. The CSF Aβ cut-off for high/ low burden ranges between 192 pg/ml and 550 pg/ml. Efforts are being made to standardize such metrics for future research [35, 36].

It is important to point out the limitations of this study. Firstly, there are differences in sample collection methods, assays and data processing pipelines between ADNI 1, ADNI 2, EDAR, and DESCRIPA. In particular, the GWAS platforms used differ between the studies. The models in this study are trained in ADNI 1, which uses a slightly older Illumina chip (Human610-Quad) than ADNI 2, EDAR, and DESCRIPA. Although this means that the data from the other cohorts may not be fully utilized, it is unlikely to cause a lack of replicability. However, it is possible that some replicability is lost due to differences in sample collection methods and data processing. Furthermore, the use of the older SNP chip (Human610-Quad in ADNI 1) may have lead to reduced SNP content and sub-optimal imputation within the ADNI 1 cohort. It is also of note that the ADNI study was included in IGAP, effect sizes from which were used to create the PGRS. We believe that the benefit of a larger sample size for training outweighs any negative impact. Despite the use of well-characterized cohorts in this study, we must point out that the sample size considered is still relatively small. In future work, this could be addressed by the use of longitudinal aging studies instead of the case/control cohort studies used here.

Finally, we have shown that the multi-modal approach used in the sub-study investigating plasma tau, did not improve predictive ability above the basic model. We used a simple additive model and more complex methods such as OmicKriging may be useful in this setting [37]. Furthermore, the standard for measuring AD pathology, in particular Aβ, is through

PET imaging. Generally, PET imaging and CSF measurements are used interchangeably but any results should be replicated using imaging based outcomes. In combination with a pathological endpoint more closely related to an AD phenotype, such as the total CSF burden used here, the suggestions presented in this discussion could improve the predictive ability of proposed markers of Aβ and tau.

CONCLUSIONS

This study has used data from the EDAR, DESCRIPA, and ADNI cohorts to investigate blood markers of Aβ and tau. We see that a case/control PGRS is no more predictive of pathology than age, gender, and APOE genotype. A sub-study shows that model accuracy is not improved by the addition of plasma tau measurements. These results emphasize that the search for predictive factors of Aβ and tau, above and beyond demographic and APOE information, is still ongoing. Better understanding of AD risk alleles, development of more sensitive assays, and studies of larger sample size are three avenues that may provide such factors. However, the clinical utility of possible predictors of brain Aβ and tau must also be investigated.

ACKNOWLEDGMENTS

This work represents independent research part funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. This study was supported by researchers at the National Institute for Health Research University College London Hospitals Biomedical Research Centre. The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement number 115372, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution. Steven Kiddle is supported by an MRC Career Development Award in Biostatistics (MR/L011859/1). Nicola Voyle is funded by the Alzheimer's Society. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

This work was also supported by awards to establish the Farr Institute of Health Informatics Research at UCL partners, London, from the Medical Research Council, Arthritis Research UK, British Heart Foundation, Cancer Research UK, Chief Scientist Office, Economic and Social Research Council, Engineering and Physical Sciences Research Council, National Institute for Health Research, National Institute for Social Care and Health Research, and Wellcome Trust (grant MR/K006584/1).

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimers Association; Alzheimers Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research and Development, LLC.; John-son and Johnson Pharmaceutical Research and Development LLC.; Lumosity; Lundbeck; Merck and Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (http://www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Authors' disclosures available online (http://j-alz.com/manuscript-disclosures/16-0707r1).

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: http://dx.doi.org/10.3233/JAD-160707.

REFERENCES

[1] Dubois B, Feldman HH, Jacova C, Hampel H, Molinuevo JL, Blennow K, Dekosky ST, Gauthier S, Selkoe D, Bateman R, Cappa S, Crutch S, Engelborghs S, Frisoni GB, Fox NC, Galasko D, Habert MO, Jicha G a Nordberg A, Pasquier F, Rabinovici G, Robert P, Rowe C, Salloway S, Sarazin M, Epelbaum S, de Souza LC, Vellas B, Visser PJ, Schneider L, Stern Y, Scheltens P, Cummings JL (2014) Advancing research diagnostic criteria for Alzheimer's disease: The IWG-2 criteria. *Lancet Neurol* **13**, 614-629.

[2] Hardy J, Higgins G (1992) Alzheimer's disease: The amyloid cascade hypothesis. *Science* **256**, 184-185.

[3] Jack CR, Knopman DS, Jagust WJ, Petersen RC, Weiner MW, Aisen PS, Shaw LM, Vemuri P, Wiste HJ, Weigand SD, Lesnick TG, Pankratz VS, Donohue MC, Trojanowski JQ (2013) Tracking pathophysiological processes in Alzheimer's disease: An updated hypothetical model of dynamic biomarkers. *Lancet Neurol* **12**, 207-216.

[4] Johnson KA, Minoshima S, Bohnen NI, Donohoe KJ, Foster NL, Herscovitch P, Karlawish JH, Rowe CC, Carrillo MC, Hartley DM, Hedrick S, Pappas V, Thies WH (2013) Appropriate use criteria for amyloid PET: A report of the Amyloid Imaging Task Force the Society of Nuclear Medicine and Molecular Imaging and the Alzheimer's Association. *Alzheimers Dement* **9**, E1-E16.

[5] Buck AK, Herrmann K, Stargardt T, Dechow T, Krause BJ, Schreyögg J (2010) Economic evaluation of PET and PET/CT in oncology: Evidence and methodologic approaches. *J Nucl Med Technol* **38**, 6-17.

[6] Menéndez-González M (2014) Routine lumbar puncture for the early diagnosis of Alzheimer's disease: Is it safe? *Front Aging Neurosci* **6**, 1-2.

[7] Bazenet C, Lovestone S (2012) Plasma biomarkers for Alzheimer's disease: Much needed but tough to find. *Biomark Med* **6**, 441-454.

[8] Cruchaga C, Ebbert MTW, Kauwe JSK (2014) Genetic discoveries in AD using CSF amyloid and tau. *Curr Genet Med Rep* **2**, 23-29.

[9] Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, Jun G, DeStefano AL, Bis JC, Beecham GW, Grenier-Boley B, Russo G, Thornton-Wells TA, Jones N, Smith AV, Chouraki V, Thomas C, Ikram MA, Zelenika D, Vardarajan BN, Kamatani Y, Lin CF, Gerrish A, Schmidt H, Kunkle B, Dunstan ML, Ruiz A, Bihoreau MT, Choi SH, Reitz C, Pasquier F, Hollingworth P, Ramirez A, Hanon O, Fitzpatrick AL, Buxbaum JD, Campion D, Crane PK, Baldwin C, Becker T, Gudnason V, Cruchaga C, Craig D, Amin N, Berr C, Lopez OL, De Jager PL, Deramecourt V, Johnston JA, Evans D, Lovestone S, Letenneur L, Moron FJ, Rubinsztein DC, Eiriksdottir G, Sleegers K, Goate AM, Fiévet N, Huentelman MJ, Gill M, Brown K, Kamboh MI, Keller L, Barberger-Gateau P, McGuinness B, Larson EB, Green R, Myers AJ, Dufouil C, Todd S, Wallon D, Love S, Rogaeva E, Gallacher J, George-Hyslop PS, Clarimon J, Lleo A, Bayer A, Tsuang DW, Yu L, Tsolaki M, Bossù, P, Spalletta G, Proitsi P, Collinge J, Sorbi S, Sanchez-Garcia F, Fox NC, Hardy J, Naranjo MCD, Bosco P, Clarke R, Brayne C, Galimberti D, Mancuso M, Matthews F, Moebus S, Mecocci P, Del Zompo M, Maier W, Hampel H, Pilotto A, Bullido M, Panza F, Caffarra P, Nacmias B, Gilbert JR, Mayhaus M, Lannfelt L, Hakonarson H, Pichler S, Carrasquillo MM, Ingelsson M, Beekly D, Alvarez V, Zou F, Valladares O, Younkin SG, Coto E, Hamilton-Nelson KL, Gu W, Razquin C, Pastor P, Mateo I, Owen MJ, Faber KM, Jonsson PV, Combarros O, O'Donovan MC, Cantwell LB, Soininen H, Blacker D, Mead S, Mosley TH, Bennett DA, Harris TB, Fratiglioni L, Holmes C, de Bruijn RFAG, Passmore P, Montine TJ, Bettens K, Rotter JI, Brice A, Morgan K, Foroud TM, Kukull WA, Hannequin D, Powell J F, Nalls MA, Ritchie K, Lunetta KL, Kauwe JSK, Boerwinkle E, Riemenschneider M, Boada M, Hiltunen M, Martin ER, Schmidt R, Rujescu D, Wang LS, Dartigues JF, Mayeux R, Tzourio C, Hofman A, Nöthen MM, Graff C, Psaty BM, Jones L, Haines JL, Holmans PA, Lathrop M, Pericak-Vance MA, Launer LJ, Farrer LA, van Duijn CM, Van Broeckhoven C, Moskvina V, Seshadri S, Williams J, Schellenberg GD, Amouyel P (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* **45**, 1452-1458.

[10] Lee SH, Harold D, Nyholt DR, ANZGene Consortium, International Endogene Consortium, Genetic and Environmental Risk for Alzheimer's disease Consortium, Goddard ME, Zondervan KT, Williams J, Montgomery GW, Wray NR, Visscher PM (2013) Estimation and partitioning of polygenic variation captured by common SNPs for Alzheimer's disease multiple sclerosis and endometriosis. *Hum Mol Genet* **22**, 832-841.

[11] Ridge PG, Mukherjee S, Crane PK, Kauwe JSK, Alzheimer's Disease Genetics Consortium (2013) Alzheimer's disease: Analyzing the missing heritability. *PLoS One* **8**, e79771.

[12] Ruiz A, Heilmann S, Becker T, Hernández I, Wagner H, Thelen M, Mauléon A, Bellenguez C (2014) Follow-up of loci from the International Genomics of Alzheimer's Disease Project identifies TRIP4 as a novel susceptibility gene. *Transl Psychiatry* **4**, e358.

[13] Ridge PG, Hoyt KB, Boehme K, Mukherjee S, Crane PK, Haines JL, Mayeux R, Farrer LA, Pericak-Vance MA, Schellenberg GD, Kauwe JSK (2016) Assessment of the genetic variance of late-onset Alzheimer's disease. *Neurobiol Aging* **41**, 200e13-200e20.

[14] Gatz M, Pedersen NL, Berg S, Johansson B, Johansson K, Mortimer J, a Posner SF, Viitanen M, Winblad B, Ahlbom A (1997) Heritability for Alzheimer's disease: The study of dementia in Swedish twins. *J Gerontol A Biol Sci Med Sci* **52A**, M117-M125.

[15] Pedersen NL, Gatz M, Berg S, Johansson B (2004) How heritable is Alzheimer's disease late in life? Findings from Swedish twins. *Ann Neurol* **55**, 180-185.

[16] Dudbridge F (2013) Power and predictive accuracy of polygenic risk scores. *PLoS Genet* **9**, e1003348.

[17] Cruchaga C, Kauwe JSK, Harari O, Jin SC, Cai Y, Karch CM, Benitez BA, Jeng AT, Skorupa T, Carrell D, Bertelsen S, Bailey M, McKean D, Shulman JM, De Jager PL, Chibnik L, Bennett DA, Arnold SE, Harold D, Sims R, Gerrish A, Williams J, Van Deerlin VM, Lee VMY, Shaw LM, Trojanowski JQ, Haines JL, Mayeux R, Pericak-Vance MA, Farrer LA, Schellenberg GD, Peskind ER, Galasko D, Fagan AM, Holtzman DM, Morris JC, Goate AM, GERAD Consortium, Alzheimer's Disease Neuroimaging Initiative, Alzheimer's Disease Genetic Consortium (2013) GWAS of cerebrospinal fluid tau levels identifies risk variants for Alzheimer's disease. *Neuron* **78**, 256-268.

[18] Sabuncu MR, Buckner RL, Smoller JW, Lee PH, Fischl B, Sperling RA, Alzheimer's Disease Neuroimaging Initiative (2012) The Association between a polygenic Alzheimer score and cortical thickness in clinically normal subjects. *Cereb Cortex* **22**, 2653-2661.

[19] Sleegers K, Bettens K, De Roeck A, Van Cauwenberghe C, Cuyvers E, Verheijen J, Struyfs H, Van Dongen J, Vermeulen

S, Engelborghs S, Vandenbulcke M, Vandenberghe R, De Deyn PP, Van Broeckhoven C, BELNEU Consortium (2015) A 22-single nucleotide polymorphism Alzheimer risk score correlates with family history onset age and cerebrospinal fluid Aβ42. *Alzheimers Dement* **11**, 1452-1460.

[20] Visser PJ, Verhey FRJ, Boada M, Bullock R, De Deyn PP, Frisoni GB, Fr ouml lich L, Hampel H, Jolles J, Jones R, Minthon L, Nobili F, Olde Rikkert M, Ousset PJ, Rigaud AS, Scheltens P, Soininen H, Spiru L, Touchon J, Tsolaki M, Vellas B, Wahlund LO, Wilcock G, Winblad B (2008) Development of Screening Guidelines and Clinical Criteria for Predementia Alzheimer's Disease. *Neuroepidemiology* **30**, 254-265.

[21] Barnett JH, Blackwell A, Scheltens P, Waldemar G, Johannsen P, Tsolaki M, Vandenberghe R, Wahlund LO, Verhey F, Visser PJ (2010) Cognitive function and cognitive change in dementia mild cognitive impairment and healthy aging: The EDAR study. *Alzheimers Dement* **6**, S127.

[22] Visser PJ, Verhey F, Knol DL, Scheltens P, Wahlund LO, Freund-Levi Y, Tsolaki M, Minthon L, Wallin ÅK, Hampel H, Bürger K, Pirttila T, Soininen H, Rikkert MO, Verbeek MM, Spiru L, Blennow K (2009) Prevalence and prognostic value of CSF markers of Alzheimer's disease pathology in patients with subjective cognitive impairment or mild cognitive impairment in the DESCRIPA study: A prospective cohort study. *Lancet Neurol* **8**, 619-627.

[23] Saykin AJ, Shen L, Foroud TM, Potkin SG, Swaminathan S, Kim S, Risacher SL, Nho K, Huentelman MJ, Craig DW, Thompson PM, Stein JL, Moore JH, Farrer LA, Green RC, Bertram L, Jack Jr CR, Weiner MW (2010) Alzheimer's Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: Genetics core aims progress and plans. *Alzheimers Dement* **6**, 265-273.

[24] Coleman JRI, Euesden J, Patel H, Folarin AA, Newhouse S, Breen G (2015) Quality control imputation and analysis of genome-wide genotyping data from the Illumina HumanCoreExome microarray. *Brief Funct Genomics* **15**, 298-304.

[25] Auton A, Abecasis GR, Altshuler DM, Durbin RM, Bentley DR, Chakravarti A, Clark AG, Donnelly P, Eichler EE, Flicek P, Gabriel SB, Gibbs RA, Green ED, Hurles ME, Knoppers BM, Korbel JO, Lander ES, Lee C, Lehrach H, Mardis ER, Marth GT, McVean GA, Nickerson DA, Schmidt JP, Sherry ST, Wang J, Wilson RK, Boerwinkle E, Doddapaneni H, Han Y, Korchina V, Kovar C, Lee S, Muzny D, Reid JG, Zhu Y, Chang Y, Feng Q, Fang X, Guo X, Jian M, Jiang H, Jin X, Lan T, Li G, Li J, Li Y, Liu S, Liu X, Lu Y, Ma X, Tang M, Wang B, Wang G, Wu H, Wu R, Xu X, Yin Y, Zhang D, Zhang W, Zhao J, Zhao M, Zheng X, Gupta N, Gharani N, Toji LH, Gerry NP, Resch AM, Barker J, Clarke L, Gil L, Hunt SE, Kelman G, Kulesha E, Leinonen R, McLaren WM, Radhakrishnan R, Roa A, Smirnov D, Smith RE, Streeter I, Thormann A, Toneva I, Vaughan B, Zheng-Bradley X, Grocock R, Humphray S, James T, Kingsbury Z, Sudbrak R, Albrecht MW, Amstislavskiy VS, Borodina TA, Lienhard M, Mertes F, Sultan M, Timmermann B, Yaspo ML, Fulton L, Fulton R, Ananiev V, Belaia Z, Beloslyudtsev D, Bouk N, Chen C, Church D, Cohen R, Cook C, Garner J, Hefferon T, Kimelman M, Liu C, Lopez J, Meric P, O'Sullivan C, Ostapchuk Y, Phan L, Ponomarov S, Schneider V, Shekhtman E, Sirotkin K, Slotta D, Zhang H, Balasubramaniam S, Burton J, Danecek P, Keane TM, Kolb-Kokocinski A, McCarthy S, Stalker J, Quail M, Davies CJ, Gollub J, Webster T, Wong B, Zhan Y, Campbell CL, Kong Y, Marcketta A, Yu F, Antunes L, Bainbridge M, Sabo A, Huang Z, Coin LJM, Fang L, Li Q, Li Z, Lin H, Liu B, Luo R, Shao H, Xie Y, Ye C, Yu C, Zhang F, Zheng H, Zhu H, Alkan C, Dal E, Kahveci F, Garrison EP, Kural D, Lee WP, Fung Leong W, Stromberg M, Ward AN, Wu J, Zhang M, Daly MJ, DePristo MA, Handsaker RE, Banks E, Bhatia G, del Angel G, Genovese G, Li H, Kashin S, McCarroll SA, Nemesh JC, Poplin RE, Yoon SC, Lihm J, Makarov V, Gottipati S, Keinan A, Rodriguez-Flores JL, Rausch T, Fritz MH, Stütz AM, Beal K, Datta A, Herrero J, Ritchie GRS, Zerbino D, Sabeti PC, Shlyakhter I, Schaffner SF, Vitti J, Cooper DN, Ball EV, Stenson PD, Barnes B, Bauer M, Keira Cheetham R, Cox A, Eberle M, Kahn S, Murray L, Peden J, Shaw R, Kenny EE, Batzer MA, Konkel MK, Walker JA, MacArthur DG, Lek M, Herwig R, Ding L, Koboldt DC, Larson D, Ye K, Gravel S, Swaroop A, Chew E, Lappalainen T, Erlich Y, Gymrek M, Frederick Willems T, Simpson JT, Shriver MD, Rosenfeld JA, Bustamante CD, Montgomery SB, De La Vega FM, Byrnes JK, Carroll AW, DeGorter MK, Lacroute P, Maples BK, Martin AR, Moreno-Estrada A, Shringarpure SS, Zakharia F, Halperin E, Baran Y, Cerveira E, Hwang J, Malhotra A, Plewczynski D, Radew K, Romanovitch M, Zhang C, Hyland FCL, Craig DW, Christoforides A, Homer N, Izatt T, Kurdoglu AA, Sinari SA, Squire K, Xiao C, Sebat J, Antaki D, Gujral M, Noor A, Ye K, Burchard EG, Hernandez RD, Gignoux CR, Haussler D, Katzman SJ, James Kent W, Howie B, Ruiz-Linares A, Dermitzakis ET, Devine SE, Min Kang H, Kidd JM, Blackwell T, Caron S, Chen W, Emery S, Fritsche L, Fuchsberger C, Jun G, Li B, Lyons R, Scheller C, Sidore C, Song S, Sliwerska E, Taliun D, Tan A, Welch R, Kate Wing M, Zhan X, Awadalla P, Hodgkinson A, Li Y, Shi X, Quitadamo A, Lunter G, Marchini JL, Myers S, Churchhouse C, Delaneau O, Gupta-Hinch A, Kretzschmar W, Iqbal Z, Mathieson I, Menelaou A, Rimmer A, Xifara DK, Oleksyk TK, Fu Y, Liu X, Xiong M, Jorde L, Witherspoon D, Xing J, Browning BL, Browning SR, Hormozdiari F, Sudmant PH, Khurana E, Tyler-Smith C, Albers CA, Ayub Q, Chen Y, Colonna V, Jostins L, Walter K, Xue Y, Gerstein MB, Abyzov A, Balasubramanian S, Chen J, Clarke D, Fu Y, Harmanci AO, Jin M, Lee D, Liu J, Jasmine Mu X, Zhang J, Zhang Y, Hartl C, Shakir K, Degenhardt J, Meiers S, Raeder B, Paolo Casale F, Stegle O, Lameijer EW, Hall I, Bafna V, Michaelson J, Gardner EJ, Mills RE, Dayama G, Chen K, Fan X, Chong Z, Chen T, Chaisson MJ, Huddleston J, Malig M, Nelson BJ, Parrish NF, Blackburne B, Lindsay SJ, Ning Z, Zhang Y, Lam H, Sisu C, Challis D, Evani US, Lu J, Nagaswamy U, Yu J, Li W, Habegger L, Yu H, Cunningham F, Dunham I, Lage K, Berg Jespersen J, Horn H, Kim D, Desalle R, Narechania A, Wilson Sayres MA, Mendez FL, David Poznik G, Underhill PA, Coin L, Mittelman D, Banerjee R, Cerezo M, Fitzgerald TW, Louzada S, Massaia A, Ritchie GR, Yang F, Kalra D, Hale W, Dan X, Barnes KC, Beiswanger C, Cai H, Cao H, Henn B, Jones D, Kaye JS, Kent A, Kerasidou A, Mathias R, Ossorio PN, Parker M, Rotimi CN, Royal CD, Sandoval K, Su Y, Tian Z, Tishkoff S, Via M, Wang Y, Yang H, Yang L, Zhu J, Bodmer W, Bedoya G, Cai Z, Gao Y, Chu J, Peltonen L, Garcia-Montero A, Orfao A, Dutil J, Martinez-Cruzado JC, Mathias RA, Hennis A, Watson H, McKenzie C, Qadri F, LaRocque R, Deng X, Asogun D, Folarin O, Happi C, Omoniwa O, Stremlau M, Tariyal R, Jallow M, Sisay Joof F, Corrah T, Rockett K, Kwiatkowski D, Kooner J, Tnh Hiê'n T, Dunstan SJ, Thuy Hang N, Fonnie R, Garry R, Kanneh L, Moses L, Schieffelin J, Grant DS, Gallo C, Poletti G, Saleheen D, Rasheed A, Brooks LD, Felsenfeld AL, McEwen JE, Vaydylevich Y, Duncanson A,

Dunn M, Schloss JA (2015) A global reference for human genetic variation. *Nature* **526**, 68-74.

[26] R Core Team (2014) R: A language and environment for statistical computing.

[27] Kuhn M, Wing J, Weston S, Williams A, Keefer C, Engelhardt A, Cooper T, Mayer Z, Kenkel B, R Core Team, Benesty M, Lescarbeau R, Ziem A, Scrucca L (2015) caret: Classification and Regression Training. R Package Version 6.0-41, http://CRAN.R-project.org/package=caret

[28] Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Müller M (2011) pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* **12**, 77.

[29] Sing T, Sander O, Beerenwinkel N, Lengauer T (2005) ROCR: Visualizing classifier performance in R. *Bioinformatics* **21**, 3940-3941.

[30] Euesden J, Lewis CM, O'Reilly PF (2015) PRSice: Polygenic risk score software. *Bioinformatics* **31**, 1466-1468.

[31] Escott-Price V, Sims R, Bannister C, Harold D, Vronskaya M, Majounie E, Badarinarayan N, Morgan K, Passmore P, Holmes C, Powell J, Brayne C, Gill M, Mead S, Goate A, Cruchaga C, Lambert J C, Van Duijn C, Maier W, Ramirez A, Holmans P, Jones L, Hardy J, Seshadri S, Schellenberg GD, Amouyel P, Williams J (2015) Common polygenic variation enhances risk prediction for Alzheimer's disease. *Brain* **138**, 3673-3684.

[32] Baird AL, Westwood S, Lovestone S (2015) Blood-based proteomic biomarkers of Alzheimer's disease pathology. *Front Neurol* **6**, 236.

[33] Voyle N, Kim M, Proitsi P, Ashton NJ, Baird AL, Bazenet C, Hye A, Westwood S, Chung R, Ward M, Rabinovici GD, Lovestone S, Breen G, Legido-Quigley C, Dobson RJB, Kiddle SJ (2016) Blood metabolite markers of neocortical amyloid-β burden: Discovery and enrichment using candidate proteins. *Transl Psychiatry* **6**, e719.

[34] Mattsson N, Zetterberg H, Blennow K (2010) Lessons from multicenter studies on CSF biomarkers for Alzheimer's disease. *Int J Alzheimers Dis* **2010**, pii:610613.

[35] García Barrado L, Coart E, Vanderstichele HM, Burzykowski T (2015) Transferring cut-off values between assays for cerebrospinal fluid Alzheimer's disease biomarkers. *J Alzheimers Dis* **49**, 187-199.

[36] Zwan MD, Rinne JO, Hasselbalch SG, Nordberg A, Lleo A, Herukka SK, Soininen H, Law I, Bahl JMC, Carter SF, Fortea J, Blesa R, Teunissen CE, Bouwman FH, van Berckel BNM, Visser PJ (2015) Use of amyloid-PET to determine cutpoints for CSF markers: A multicenter study. *Am Acad Neurol* **86**, 50-58.

[37] Wheeler HE, Aquino-Michaels K, Gamazon ER, Trubetskoy VV, Dolan ME, Huang RS, Cox NJ, Im HK (2014) Poly-omic prediction of complex traits: OmicKriging. *Genet Epidemiol* **38**, 402-415.

Chapter 7

Discussion

7.1 Summary of findings

In this thesis I aimed to identify blood-based biomarkers of AD and it's associated phenotypes, in particular Aβ and tau burden. I have studied a variety of components in blood including genetics, gene expression, proteomics and metabolomics using data from several cohort studies to explore single and multi-modal biomarkers of AD.

7.1.1 Single modality markers

7.1.1.1 Chapter 2: Blood protein markers of neocortical amyloid-β burden

In the search for a blood biomarker of AD, replication of existing discoveries was an essential first step. Forty-one proteomic markers of Aβ were identified from the literature: 15 candidates discovered in AIBL and 20 from other cohorts. Two of these candidates (IgM and PPY) met a pre-defined significance threshold and were hence deemed replicable in our SOMAscan data from AIBL subjects. This replication adds evidence to the association of IgM and PPY with brain amyloid burden despite technical differences between studies, such as biological fluid and platform. However, it is possible that the lack of replication of other candidates was also due to these technical differences.

7.1.1.2 Chapter 3: A pathway based classification method for analyzing gene expression for AD diagnosis

In chapter 3, I hypothesised that combining gene expression levels to create a pathway level score may improve the ability of a model to differentiate between AD cases and controls. The work in AddNeuroMed concluded that the pathway level models were no better than gene expression or indeed demographic only models in the available data. Independent validation of this work is essential. This is the only chapter of this thesis that discusses a case/control endpoint and an endophenotype based approach may be beneficial for future work in this area.

7.1.2 Multiple modality markers

7.1.2.1 Chapter 4: Blood metabolite markers of neocortical amyloid-β burden

The motivation for studying metabolites as blood markers of AD is driven by their intuitive link with the disease. In particular, the fact that a major risk gene for AD, *APOE*, codes for a protein that transports fats and fat soluble vitamins which are both metabolites. This chapter should be viewed as a pilot study due to the small number of subjects. It provides encouraging evidence for associations between metabolites and amyloid which warrant further investigation and attempts at replication. Furthermore, this chapter provides evidence for the potential predictive ability of multi-modal biomarkers, exemplified by the addition of protein measurements to the metabolite model.

7.1.2.2 Chapter 5: Do peripheral markers help to predict amyloid burden in a non-demented population? A Bayesian approach.

In Chapter 5, I present a Bayesian methodology that aims to incorporate historical information on demographic variables into the prediction of amyloid in a non-demented population. The hypothesis that this method would improve predictive ability was not shown by the results. Further, the focus of the study was using demographics and a PGRS to predict CSF amyloid. The inclusion of the PGRS

did not improve predictive ability over demographics alone. Furthermore, I hypothesised that inclusion of multiple modalities of *omics* data in modelling would improve predictive ability; this was not shown by the results of this analysis.

7.1.2.3 Chapter 6: Genetic risk as a marker of amyloid-β and tau burden in cerebrospinal fluid.

The final chapter of this thesis aims to examine a PGRS as a marker of tau levels in CSF. Additionally, I investigate the PGRS as a marker of a combined amyloid and tau endpoint representing a more 'AD-like' phenotype. In the populations studied, the PGRS was not significantly associated with AD pathologies and did not improve predictive ability over demographics. Further, I performed a sub-study with the inclusion of plasma tau as a predictor. Similarly, no improvements in predictive ability were seen.

7.2 Implications of findings

7.2.1 Clinical

Due to the lack of replication of the biomarkers presented in this thesis the current clinical implications are minimal. For a biomarker of AD or pathology burden to have clinical utility it would need to be replicable across independent cohorts. This level of reproducibility and consistency has not yet been shown for any biomarkers discussed in this thesis. Nevertheless, the replication of two candidate proteins and discovery of a 5-metabolite panel whose predictive ability is improved by the addition of FGG, indicates that these methods warrant further investigation and could lead to clinical utility in the future.

Additionally, the results outlined in this thesis could provide insight into possible disease mechanisms of AD or Aβ accumulation. Conventionally, to find biomarkers of a disease one would target compounds that are thought (or known) to be involved in the causal process behind the disease but in AD no such process is known. Therefore, we can look to inform existing theories using the associations shown in

this thesis. It is of course vital to ensure that these associations are not interpreted as causal mechanisms. With this in mind it is equally important for negative results, as well as positive findings, to be communicated to clinicians. Such dissemination ensures all evidence is available to key decision makers.

7.2.2 Research

In comparison to the clinical implications the impact for research is more obvious. Firstly, I have shown that associations of some proteins (IgM and PPY) with Aβ seem to replicate across technical platforms and studies. Of particular interest is the association of IgM with Aβ burden in a non-demented population showing promise for use as a filtering step for enrichment in secondary prevention clinical trials. Furthermore, I have shown associations of a metabolite panel with Aβ burden in a largely demented population. This work was performed in a relatively small number of subjects and provides a basis for further research. Others have shown replication of protein markers between blood fractions and using different proteomic platforms (O'Bryant & Xiao, 2014). This method of systematic alteration of technical factors builds the evidence for association of specific markers with AD phenotypes. Further, we can begin to address which set of technical factors gives rise to the best biomarker. For example, O'Bryant & Xiao (2014) find that markers in blood serum are more predictive of AD cases than equivalent markers in blood plasma.

Secondly, this thesis provides some evidence for the use of a multi-modality approach when using *omics* data in the search for a biomarker of Aβ. Chapter 3 is the first published piece of work to exemplify this although the theoretical idea has been developing for some time (Bazenet & Lovestone, 2012; Snyder *et al.*, 2014). Chapters 5 and 6 challenge this idea as the addition of multiple modalities of *omics* data did not lead to improvement in the predictive ability of models. As highlighted in the individual chapters this could be due to the PGRS being trained on a case/control endpoint or the selection of candidate protein and metabolite markers. As more cohorts with a variety of *omics* data become available this avenue will hold lots of research potential. For example, the European Medical Information

Framework (EMIF) 1000 core is a dataset that aims to have PET scanning and blood *omics* measurements available for 1000 individuals by the end of 2016 through EMIF funding. Furthermore, the development of methods to more effectively combine the independent signal seen in each modality will be vital.

More generally, the work in this thesis highlights the importance of replication in all discovery work, not just within *omics* analysis or AD. Through Chapters 5 and 6 I have shown that model metrics (in particular accuracy, sensitivity and specificity) decrease substantially when models are tested on independent datasets with previously un-modelled noise. This was also exemplified by Casanova *et al.* (2016) who failed to replicate a metabolite signature of AD published by Mapstone *et al.* (2014).

7.3 Limitations

This section details some research limitations that affect the whole thesis. More specific limitations related to each analysis are given in individual chapters.

A limitation that has recurred throughout this work is a lack of statistical power largely driven by a small number of participants in each analysis. This is particularly true of Chapter 4 (N = 78, metabolic features = 2760). In all analyses I have aimed to maximise sample sizes where possible, for example by using the measurement of Aβ with largest sample size available. Furthermore, the Bayesian method employed in Chapter 5 was motivated by the relatively small size of current AD *omics* studies. Lower participant numbers are largely due to a small overlap in the people with amyloid measurements and *omics* data. Additionally, in the gene expression analysis I chose to study as clean a phenotype as possible by only including cases and controls that had maintained diagnosis at all visits. Overall, larger studies are needed to ensure any significant associations are real and to minimise the number of false negative associations. It is becoming more likely that these larger sample sizes will originate from population studies of health cohorts rather than cohort studies of individuals with AD and other dementias. For example, the PROTECT study (www.protectstudy.org.uk) is recruiting individuals over the age of 50 with no diag-

nosis of dementia (N $\sim$ 5000). It aims to follow them for 10 years to elucidate factors that influence changes in the brain during ageing. A sub-group of individuals will have blood measurements collected and will provide an interesting asymptomatic discovery population.

The power available in AD cohort studies is also limited by the presence of mixed pathologies. Although I have focused on modelling associations with Aβ (and in some cases tau) pathology we cannot be certain that there are not other confounding factors such as presence of other forms of dementia. For example, in the UCSF cohort used in Chapter 4 a large number of individuals were diagnosed with fronto-temporal dementia. Furthermore, it is possible that biomarkers investigated in this thesis are confounded by medications the participants are taking. This was difficult to investigate in several of the cohorts used in this thesis as medication information was not available. However, Tao *et al.* (2015) have shown no significant confounding of treatment for several gene expression and protein biomarkers. I am a co-author on this paper for providing the processed gene expression data generated in chapter 3 of this thesis (Appendix A).

Across the field of AD research there is a lack of knowledge of the functional role of amyloid in ageing and AD. In particular, there are distinct examples of Aβ not always causing dementia symptoms, with people who have high amyloid burden at autopsy having shown no cognitive impairment during their lifetime (Morris *et al.*, 2014). Current theories hypothesise that this could be because they have not yet accumulated enough pathology or because factors such as cognitive reserve influence the effect of any pathology. This raises the question that even if we can predict Aβ burden using a biomarker, will anti-amyloid treatments necessarily lead to improvements in cognition? At least one clinical trial to date saw improvements in Aβ levels measured by PET imaging but no significant changes in cognitive symptoms (Liu *et al.*, 2015). This lack of improvement could be explained by the fact that PET imaging measures the plaque formations of Aβ but not the oligomeric amyloid which is thought to be the toxic form. If methods are developed to measure this oligomeric form answers to this question could be found through further studies.

Finally, across science there is a tendency to publish positive results more often than negative findings, leading to publication bias (Song *et al.*, 2010). This can mean that a sample of studies is unrepresentative of all analysis that has been performed in the research area. One solution to prevent publication bias is to enforce (or at least recommend) registration of all studies to some central database. As such, several individuals have called for detailed registries of biomarker studies to be established (Andre *et al.*, 2011; Altman, 2014). If implemented, these registries would better inform evidence based medicine in the field of AD biomarker research.

7.4 Further work

With the research implications outlined above in mind there are several avenues that should be investigated further.

Firstly, work in this thesis has demonstrated that multi-modal analysis is possible (Chapters 4, 5 and 6). However, it has not always improved the predictive ability of a model and in all cases simple additive modelling techniques have been used. It would be interesting to develop more sophisticated methods that could ensure any information included by an additional modality is independent and hence potentially useful. Furthermore, interactions between modalities should be investigated. The work in this thesis has not implemented feature selection across modalities. Without these developments there is a risk of creating models that over-fit to the training data with minimal additional benefit. Furthermore, economic cost-benefit analyses will become even more important at this stage. We need to ensure that any potential blood-test that is developed will be more cost-effective than the current standard for measuring Aβ and tau (a PET scan or lumbar puncture). As we add more modalities to our modelling this becomes less likely.

Chapter 4 presents the first study to investigate associations between Aβ and metabolites. I was able to putatively identify the metabolites of interest through comparisons with the Human Metabolome Database and in-house tables. To be more confident in this identification we could purchase standards of these metabolites (where available) and compare their LC-MS/MS spectra with that seen in

our study. This non-putative identification would allow targeted analysis to be performed using a variety of platforms in the future. Furthermore, the levels of metabolites in each individual were calculated using the R package 'XCMS'. By performing semi-quantification (measuring the area under the chromatogram peak by hand) for the metabolic features of interest, we could be more certain of their levels and hence the associations we have seen in Chapter 4. The metabolite study is the only study in this thesis to contain a population that predominantly contains people with AD or FTD. This was well documented in the chapter itself and it provides opportunity for further study in a non-demented population; the population that could see most benefit from such a blood test being used as a filtering step. This notion is applicable to all research in this area: the early stages of disease development are vital and consequently research should focus on these populations.

In further work it would be interesting to covary for novel factors that are appearing in the literature to be associated with Aβ and tau burden. For example, it has been shown in recent studies that levels of 'brain reserve' may be driving the difference between people who have high levels of pathology and no symptoms and those who show symptoms (Bauckneht *et al.*, 2015). This is motivating a theory that increased brain reserve may prevent development of symptoms even if pathology is present. Although brain reserve itself may be hard to quantify it may be possible to measure and model associated psychological factors such as social networks.

There are also some more general research questions that if addressed would improve the validity and replicability of work in this area. Firstly, it is well documented that a wide range of cut-offs are used to differentiate individuals who have high levels of pathology from those that have low levels, particularly in CSF (Mattsson *et al.*, 2010). This is largely due to differences in assays but has also been shown to be influenced by technical differences in processing methods. A standardised protocol, and potentially assay, would alleviate this problem and specialised working groups are discussing this currently (Zwan *et al.*, 2015). This is an area that should be closely monitored by the AD biomarker research community and any recommended practise implemented as soon as possible. This lack of comparability between assays

also extends to the majority of *omics* platforms (Mattsson *et al.*, 2010). For example, Casanova *et al.* (2016) failed to replicate a set of predictive metabolite markers discovered by Mapstone *et al.* (2014). It was suggested that this failed replication was due to differences between blood plasma and serum. However, high correlations between metabolites in blood plasma and serum in other studies make this unlikely (Casanova *et al.*, 2016). Several other factors were also altered between the studies, for example processing methods. Therefore, we cannot be sure what is driving this lack of replication. If *omics* studies were standardised it would be easier to compare findings and elucidate which factors were driving any difference in results.

The majority of the work described in this thesis focuses on a single measurement of amyloid that in most cases is then dichotomised to either high or low. This is well justified by the bimodal distribution of amyloid burden across a population. However, it is well known that there are other pathologies and processes associated with development of AD. I believe an important next step is to find biomarkers of these pathologies and combinations of them with Aβ burden. I began to address this by studying possible markers of tau and a combined Aβ and tau endpoint in Chapter 6. As tau imaging tracers are now becoming available I think this would be an increasingly interesting endophenotype to study. Additionally, a distributional measure of the spread of pathology in the brain may be more informative than an overall Aβ or tau burden metric. Analysis of the spread of pathologies in the brain have shown this area of research could hold promise (Thal *et al.*, 2014).

If robust and replicable blood biomarkers are discovered they could have immediate impact in trials such as the European Prevention of Alzheimer's Dementia (EPAD) trial. EPAD is an adaptive trial meaning many potential treatments are tested against each other and against placebo at the same time. At pre-specified interim analyses researchers can determine whether certain treatments should be removed from the trial. Novel treatments can also be added. Recruitment into the EPAD trial is largely through registries from existing cohort studies. However, individuals eligible to participate must undergo brain imaging to ensure they have the amyloid pathology many of the experimental treatments will target. The use of a

blood biomarker has the potential to significantly reduce the cost of this screening step. Furthermore, blood biomarkers could be used to monitor amyloid burden at the interim analyses again, potentially reducing the cost of longitudinal monitoring in the study.

The research in this thesis, amongst other literature, highlights that there is still a long way to go in the discovery of a blood based biomarker for AD. The data analysis methods used throughout this thesis have consistently been shown to achieve good performance in other areas (Touw *et al.*, 2012). I therefore suggest that the main areas of opportunity to improve discoveries in this field come in the form of data quality and quantity. The quality of *omics* data currently available is limited by technical variability in collection and processing methods (O'Bryant *et al.*, 2015). Techniques to reduce such variability should be developed and implemented in future studies. Furthermore, even when technical variability is controlled the assays available are often sub-optimal. For example, the specificity of methods such as SOMAscan is largely unknown and sensitivity of methods including MS can be low. New methodologies such as SIMOA begin to address this by measuring quantities by counting single molecules. Further novel assay development will be vital in finding a blood based biomarker for AD.

Data quantity in cohort studies is also a limiting factor to AD blood biomarker discovery. With a large number of variables, often in the thousands, the benefit that larger numbers of individuals bring is unquestionable. Initiatives such as the EMIF core 1000 are beginning to address this. However, to make the most of this opportunity it is essential that such studies are well planned. For example, the overlap in samples with different types of data collected should be as large as possible. This will maximize the number of individuals with multiple measurements to enable the study of combinations of *omics* markers. Finally, the use of longitudinal data will provide invaluable information. The deep and frequent phenotyping study will be one of the first resources to provide such data in AD (www.psych.ox.ac.uk/research/translational-neuroimaging-group/team/deep-frequent-phenotyping).

7.5 Conclusions

Global dementia prevalence has been estimated at 24.3 million people (Ferri *et al.*, 2005). With an ageing population AD, as the most common form of dementia, poses a significant socio-economic burden to society. Evidence indicates that pathological changes leading to AD can occur up to 20 years prior to symptom development. This provides a window of opportunity to target disease-modifying treatments yet, to date, none are available. To improve the likelihood of finding a suitable treatment clinical trials are being enriched for individuals with high levels of Aβ pathology as quantified by PET imaging or measurements in CSF. These methods of measuring pathology are often sub-optimal being invasive and expensive procedures that require specialist services. Consequently, this thesis presents work searching for a blood-based biomarker for use as a filtering step prior to a confirmatory PET scan or lumbar puncture.

No biomarkers identified this far have reached clinical utility. However, I have shown that there is validity in using *omics* data to look for associations with AD pathology, in particular amyloid, in the search for a biomarker. Further, the use of a combination of markers from multiple modalities warrants more research. Work throughout this thesis demonstrates the need for replication. This will become increasingly possible in the near future as larger cohorts with a variety of *omics* data become available.

If biomarkers of AD, and in particular Aβ burden, can be identified there is great potential for cost-effective enrichment of clinical trials to aid in the discovery of a treatment for AD.

Bibliography

Altman, D.G. (2014). The Time Has Come to Register Diagnostic and Prognostic Research. *Clinical Chemistry*, **60**, 580–582.

Alzheimer's Association (2015). Alzheimer's disease facts and figures. Tech. Rep. 3, Alzheimer's Association.

Andre, F., McShane, L.M., Michiels, S., Ransohoff, D.F., Altman, D.G., Reis-Filho, J.S., Hayes, D.F. & Pusztai, L. (2011). Biomarker studies: a call for a comprehensive biomarker study registry. *Nature Publishing Group*, **8**, 171–176.

Appleby, B.S., Nacopoulos, D., Milano, N., Zhong, K. & Cummings, J.L. (2013). A Review: Treatment of Alzheimer's Disease Discovered in Repurposed Agents. *Dementia and Geriatric Cognitive Disorders*, **35**, 1–22.

Ariza, M., Kolb, H.C., Moechars, D., Rombouts, F. & Andrés, J.I. (2015). Tau Positron Emission Tomography (PET) Imaging: Past, Present, and Future. *Journal of Medicinal Chemistry*, **58**, 4365–4382.

Ashton, N.J., Kiddle, S.J., Graf, J., Ward, M., Baird, A.L., Hye, A., Westwood, S., Wong, K.V., Dobson, R.J., Rabinovici, G.D., Miller, B.L., Rosen, H.J., Torres, A., Zhang, Z., Thurfjell, L., Covin, A., Hehir, C.T., Baker, D., Bazenet, C. & Lovestone, S. (2015). Blood protein predictors of brain amyloid for enrichment in clinical trials? *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, **1**, 48–60.

Ballabh, P., Braun, A. & Nedergaard, M. (2004). The bloodbrain barrier: an overview. *Neurobiology of Disease*, **16**, 1–13.

Bauckneht, M., Picco, A., Nobili, F. & Morbelli, S. (2015). Amyloid positron emission tomography and cognitive reserve. *World Journal of Radiology*, **7**, 475–483.

Bazenet, C. & Lovestone, S. (2012). Plasma biomarkers for Alzheimer's disease: much needed but tough to find. *Biomarkers in Medicine*, **6**, 441–454.

Bien-ly, N., Boswell, C.A., Jeet, S., Devoss, J., Brug, M.V.D., Watts, R.J., Bienly, N., Boswell, C.A., Jeet, S., Beach, T.G., Hoyte, K., Luk, W. & Shihadeh, V. (2015). Lack of Widespread BBB Disruption in Alzheimer's Disease Models: Focus on Therapeutic Antibodies. *Neuron*, **88**, 289–297.

Boulesteix, A.L. & Hothorn, T. (2010). Testing the additional predictive value of high-dimensional molecular data. *BMC bioinformatics*, **11**, 78.

Breiman, L. (2001). Random forests. *Machine Learning*, **45**, 5–32.

Bulik-Sullivan, B.K., Loh, P.R., Finucane, H.K., Ripke, S., Yang, J., Patterson, N., Daly, M.J., Price, A.L. & Neale, B.M. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics*, **47**, 291–295.

Burnham, S.C., Faux, N.G., Wilson, W., Laws, S.M., Ames, D., Bedo, J., Bush, A.I., Doecke, J.D., Ellis, K.A., Head, R., Jones, G., Kiiveri, H., Martins, R.N., Rembach, A., Rowe, C.C., Salvado, O., Macaulay, S.L., Masters, C.L. & Villemagne, V.L. (2013). A blood-based predictor for neocortical Aβ burden in Alzheimer's disease: results from the AIBL study. *Molecular Psychiatry*, **19**, 519–526.

Casanova, R., Varma, S., Simpson, B., Kim, M., An, Y., Saldana, S., Riveros, C., Moscato, P., Griswold, M., Sonntag, D., Wahrheit, J., Klavins, K., Jonsson, P.V., Eiriksdottir, G., Aspelund, T., Launer, L.J., Gudnason, V., Quigley, C.L. & Thambisetty, M. (2016). Blood metabolite markers of preclinical Alzheimer's disease in two longitudinally followed cohorts of older individuals. *Alzheimer's & Dementia*, Epub ahead of print.

Citron, M. (2010). Alzheimer's disease: strategies for disease modification. *Nature Reviews Drug Discovery*, **9**, 387–398.

Clark, C.M., Pontecorvo, M.J., Beach, T.G., Bedell, B.J., Coleman, R.E., Doraiswamy, P.M., Fleisher, A.S., Reiman, E.M., Sabbagh, M.N., Sadowsky, C.H., Schneider, J.A., Arora, A., Carpenter, A.P., Flitter, M.L., Joshi, A.D., Krautkramer, M.J., Lu, M., Mintun, M.A. & Skovronsky, D.M. (2012). Cerebral PET with florbetapir compared with neuropathology at autopsy for detection of neuritic amyloid-β plaques: a prospective cohort study. *The Lancet Neurology*, **11**, 669–678.

Corbett, A. & Ballard, C. (2012). New and emerging treatments for Alzheimer's disease. *Expert Opinion on Emerging Drugs*, **17**, 147–156.

Cortes, C. & Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, **20**, 273–297.

Cruchaga, C., Kauwe, J.S.K., Harari, O., Jin, S.C., Cai, Y., Karch, C.M., Benitez, B.A., Jeng, A.T., Skorupa, T., Carrell, D., Bertelsen, S., Bailey, M., McKean, D., Shulman, J.M., De Jager, P.L., Chibnik, L., Bennett, D.A., Arnold, S.E., Harold, D., Sims, R., Gerrish, A., Williams, J., Van Deerlin, V.M., Lee, V.M.Y., Shaw, L.M., Trojanowski, J.Q., Haines, J.L., Mayeux, R., Pericak-Vance, M.A., Farrer, L.A., Schellenberg, G.D., Peskind, E.R., Galasko, D., Fagan, A.M., Holtzman, D.M., Morris, J.C., Goate, A.M., Consortium, G., 26, A.D.N.I.A. & ADGC, A.D.G.C. (2013). GWAS of Cerebrospinal Fluid Tau Levels Identifies Risk Variants for Alzheimer's Disease. *Neuron*, **78**, 256–268.

Cummings, J.L., Morstorf, T. & Zhong, K. (2014). Alzheimer's disease drug-development pipeline: few candidates, frequent failures. *Alzheimer's research & therapy*, **6**, 37.

Dubois, B., Feldman, H.H., Jacova, C., Hampel, H., Molinuevo, J.L., Blennow, K., Dekosky, S.T., Gauthier, S., Selkoe, D., Bateman, R., Cappa, S., Crutch, S., Engelborghs, S., Frisoni, G.B., Fox, N.C., Galasko, D., Habert, M.O., Jicha, G.a., Nordberg, A., Pasquier, F., Rabinovici, G., Robert, P., Rowe, C., Salloway, S., Sarazin, M., Epelbaum, S., de Souza, L.C., Vellas, B., Visser, P.J., Schneider, L., Stern, Y., Scheltens, P. & Cummings, J.L. (2014). Advancing research diagnostic

criteria for Alzheimer's disease: The IWG-2 criteria. *The Lancet Neurology*, **13**, 614–629.

EGAPP Working Group (2016). Recommendations from the EGAPP Working Group : does the use of Oncotype DX tumor gene expression profiling to guide treatment decisions improve outcomes in patients with breast cancer ? *Genetics in Medicine*, **18**, 770–779.

Ellis, B., Hye, A. & Snowden, S.G. (2015). Metabolic Modifications in Human Biofluids Suggest the Involvement of Sphingolipid, Antioxidant, and Glutamate Metabolism in Alzheimer's Disease Pathogenesis. *Journal of Alzheimer's Disease*, **46**, 313–327.

Euesden, J., Lewis, C.M. & O'Reilly, P.F. (2015). PRSice: Polygenic Risk Score software. *Bioinformatics*, **31**, 1466–1468.

Ferri, C., Prince, M., Brayne, C., Brodaty, H., Fratiglioni, L., Ganguli, M., Hall, K., Hasegawa, K., Hendrie, H., Huang, Y. & Others (2005). Global prevalence of dementia: a Delphi consensus study. *The Lancet*, **366**, 2112–2117.

Folstein, M.F., Folstein, S.E. & McHugh, P.R. (1975). Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, **12**, 189–198.

Gatz, M., Pedersen, N.L., Berg, S., Johansson, B., Johansson, K., Mortimer, J.a., Posner, S.F., Viitanen, M., Winblad, B. & Ahlbom, A. (1997). Heritability for Alzheimer's disease: the study of dementia in Swedish twins. *The journals of gerontology. Series A, Biological sciences and medical sciences*, **52A**, M117–M125.

Glenner, G.G. & Wong, C.W. (1984). Alzheimer's disease: Initial report of the purification and characterization of a novel cerebrovascular amyloid protein. *Biochemical and Biophysical Research Communications*, **120**, 885–890.

Goedert, M. (2004). Tau protein and neurodegeneration. *Seminars in Cell and Developmental Biology*, **15**, 45–49.

Guerreiro, R., Wojtas, A., Bras, J., Carrasquillo, M., Rogaeva, E., Majounie, E., Cruchaga, C., Sassi, C., Kauwe, J.S.K., Younkin, S., Hazrati, L., Collinge, J., Pocock, J., Lashley, T., Williams, J., Lambert, J.C., Amouyel, P., Goate, A., Rademakers, R., Morgan, K., Powell, J., St George-Hyslop, P., Singleton, A. & Hardy, J. (2013). TREM2 variants in Alzheimer's Disease. *New England Journal of Medicine*, **368**, 117–127.

Han, G., Wang, J., Zeng, F., Feng, X., Yu, J., Cao, H.Y., Yi, X., Zhou, H., Jin, L.W., Duan, Y., Wang, Y.J. & Lei, H. (2013). Characteristic Transformation of Blood Transcriptome in Alzheimer's Disease. *Journal of Alzheimer's Disease*, **35**, 373–386.

Hand, D., Mannila, H. & Smyth, P. (2001). *Principles of Data Mining*. MIT Press.

Hiltunen, M., van Groen, T. & Jolkkonen, J. (2009). Functional Roles of Amyloid-beta protein precursor and amyloid-beta peptides: Evidence from experimental studies. *Journal of Alzheimer's Disease*, **18**, 401–412.

Hoefer, I.E., Steffens, S., Ala-korpela, M., Ba, M., Badimon, L., Boulanger, C.M., Caligiuri, G., Dimmeler, S., Egido, J., Evans, P.C., Guzik, T., Kwak, B.R., Landmesser, U., Mayr, M., Weber, C., Pasterkamp, G. & Tun, J. (2015). Novel methodologies for biomarker discovery in atherosclerosis. *European Heart Journal*, **36**, 2635–2642.

Hu, Y., Li, L., Ehm, M.G., Bing, N., Song, K., Nelson, M.R., Talmud, P.J., Hingorani, A.D., Kumari, M., Kivimäki, M., Xu, C.F., Waterworth, D.M., Whittaker, J.C., Abecasis, G.R., Spino, C. & Kang, H.M. (2013). The Benefits of Using Genetic Information to Design Prevention Trials. *The American Journal of Human Genetics*, **92**, 547–557.

Hughes, C.P., Berg, L., Danziger, W.L., Coben, L. & Martin, R.L. (1982). A New Clinical Scale for the Staging of Dementia. *The British Journal of Psychiatry*, **140**, 566–572.

Jack, C.R., Knopman, D.S., Jagust, W.J., Petersen, R.C., Weiner, M.W., Aisen, P.S., Shaw, L.M., Vemuri, P., Wiste, H.J., Weigand, S.D., Lesnick, T.G., Pankratz, V.S., Donohue, M.C. & Trojanowski, J.Q. (2013). Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *The Lancet Neurology*, **12**, 207–216.

Jansen, W.J., Ossenkoppele, R., Knol, D.L., Tijms, B.M., Scheltens, P., Verhey, F.R.J., Visser, P.J., Aalten, P., Aarsland, D., Alcolea, D., Alexander, M., Almdahl, I.S., Arnold, S.E., Baldeiras, I., Barthel, H., van Berckel, B.N.M., Bibeau, K., Blennow, K., Brooks, D.J., van Buchem, M.A., Camus, V., Cavedo, E., Chen, K., Chételat, G., Cohen, A.D., Drzezga, A., Engelborghs, S., Fagan, A.M., Fladby, T., Fleisher, A.S., van der Flier, W.M., Ford, L., Förster, S., Fortea, J., Foskett, N., Frederiksen, K.S., Freund-Levi, Y., Frisoni, G.B., Froelich, L., Gabryelewicz, T., Gill, K.D., Gkatzima, O., Gómez-Tortosa, E., Gordon, M.F., Grimmer, T., Hampel, H., Hausner, L., Hellwig, S., Herukka, S.K., Hildebrandt, H., Ishihara, L., Ivanoiu, A., Jagust, W.J., Johannsen, P., Kandimalla, R., Kapaki, E., Klimkowicz-Mrowiec, A., Klunk, W.E., Köhler, S., Koglin, N., Kornhuber, J., Kramberger, M.G., Van Laere, K., Landau, S.M., Lee, D.Y., de Leon, M., Lisetti, V., Lleo, A., Madsen, K., Maier, W., Marcusson, J., Mattsson, N., de Mendonça, A., Meulenbroek, O., Meyer, P.T., Mintun, M.A., Mok, V., Molinuevo, J.L., Møllergård, H.M., Morris, J.C., Mroczko, B., Van der Mussele, S., Na, D.L., Newberg, A., Nordberg, A., Nordlund, A., Novak, G.P., Paraskevas, G.P., Parnetti, L., Perera, G., Peters, O., Popp, J., Prabhakar, S., Rabinovici, G.D., Ramakers, I.H.G.B., Rami, L., Resende de Oliveira, C., Rinne, J.O., Rodrigue, K.M., Rodríguez-Rodríguez, E., Roe, C.M., Rot, U., Rowe, C.C., Rüther, E., Sabri, O., Sanchez-Juan, P., Santana, I., Sarazin, M., Schröder, J., Schütte, C., Seo, S.W., Soetewey, F., Soininen, H., Spiru, L., Struyfs, H., Teunissen, C.E., Tsolaki, M., Vandenberghe, R., Verbeek, M.M., Villemagne, V.L., Vos, S.J.B., van Waalwijk van Doorn, L.J.C., Waldemar, G., Wallin, A., Wallin, Å.K., Wiltfang, J., Wolk, D.A., Zboch, M. & Zetterberg, H. (2015). Prevalence of Cerebral Amyloid Pathology in Persons Without Dementia. *JAMA*, **313**, 1915–1924.

Jellinger, K.A., Janetzky, B., Attems, J. & Kienzl, E. (2008). Biomarkers for early diagnosis of Alzheimer disease : ALZheimer ASsociated gene' a new blood biomarker? *Journal of Cellular and Molecular Medicine*, **12**, 1094–1117.

Johnson, K.A., Fox, N.C., Sperling, R.A. & Klunk, W.E. (2012). Brain imaging in Alzheimer disease. *Cold Spring Harbor perspectives in medicine*, **2**, a006213.

Johnson, K.A., Minoshima, S., Bohnen, N.I., Donohoe, K.J., Foster, N.L., Herscovitch, P., Karlawish, J.H., Rowe, C.C., Carrillo, M.C., Hartley, D.M., Hedrick, S., Pappas, V. & Thies, W.H. (2013). Appropriate use criteria for amyloid PET: A report of the Amyloid Imaging Task Force, the Society of Nuclear Medicine and Molecular Imaging, and the Alzheimer's Association. *Alzheimer's & Dementia*, **9**, E1–E16.

Jonsson, T., Stefansson, H., Steinberg, S., Jonsdottir, I., Jonsson, P.V., Snaedal, J., Bjornsson, S., Huttenlocher, J., Levey, A.I., Lah, J.J., Rujescu, D., Hampel, H., Giegling, I., Andreassen, O.A., Engedal, K., Ulstein, I., Djurovic, S., Ibrahim-Verbaas, C., Hofman, A., Ikram, M.A., van Duijn, C.M., Thorsteinsdottir, U., Kong, A. & Stefansson, K. (2013). Variant of TREM2 Associated with the Risk of Alzheimer's Disease. *New England Journal of Medicine*, **368**, 107–116.

Kaddurah-Daouk, R., Rozen, S., Matson, W., Han, X., Hulette, C.M., Burke, J.R., Doraiswamy, P.M. & Welsh-Bohmer, K.A. (2011). Metabolomic changes in autopsy-confirmed Alzheimer's disease. *Alzheimer's & Dementia*, **7**, 309–317.

Karran, E., Mercken, M. & De Strooper, B. (2011). The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics. *Nature Reviews Drug Discovery*, **10**, 698–712.

Katz, R. (2004). Biomarkers and surrogate markers: An FDA perspective. *NeuroRX*, **1**, 189–195.

Kehoe, E.G., McNulty, J.P., Mullins, P.G. & Bokde, A.L.W. (2014). Advances in MRI biomarkers for the diagnosis of Alzheimer's disease. *Biomarkers in Medicine*, **8**, 1151–1169.

Kiddle, S.J., Thambisetty, M., Simmons, A., Riddoch-Contreras, J., Hye, A., Westman, E., Pike, I., Ward, M., Johnston, C., Lupton, M.K., Lunnon, K., Soininen, H., Kloszewska, I., Tsolaki, M., Vellas, B., Mecocci, P., Lovestone, S., Newhouse, S., Dobson, R. & for the Alzheimers Disease Neuroimaging Initiative (2012). Plasma Based Markers of [11C] PiB-PET Brain Amyloid Burden. *PLOS ONE*, **7**, e44260.

Kiddle, S.J., Sattlecker, M., Proitsi, P., Simmons, A., Westman, E., Bazenet, C., Nelson, S.K., Williams, S., Hodges, A., Johnston, C., Soininen, H., K??oszewska, I., Mecocci, P., Tsolaki, M., Vellas, B., Newhouse, S., Lovestone, S. & Dobson, R.J.B. (2014). Candidate blood proteome markers of Alzheimer's disease onset and progression: A systematic review and replication study. *Journal of Alzheimer's Disease*, **38**, 515–531.

Kim, S., Swaminathan, S., Shen, L., Risacher, S.L., Nho, K., Foroud, T., Shaw, L.M., Trojanowski, J.Q., Potkin, S.G., Huentelman, M.J., Craig, D.W., DeChairo, B.M., Aisen, P.S., Petersen, R.C., Weiner, M.W., Saykin, A.J. & for ADNI (2011). Genome-wide association study of CSF biomarkers Abeta, t-tau and p-tau in the ADNI cohort. *Neurologyy*, **76**, 69–79.

King, M.B. & Rwegerera, G.M. (2014). An audit of consent practices and perceptions of lumbar puncture, Botswana inpatient setting experience. *African Journal of Emergency Medicine*, **5**, 66–69.

Kohavi, R. (1995). A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. *International Joint Conference on Artificial Intelligence*, **14**, 1137–1143.

Kraemer, S., Vaught, J.D., Bock, C., Gold, L., Katilius, E., Keeney, T.R., Kim, N., Saccomano, N.A., Wilcox, S.K., Zichi, D. & Sanders, G.M. (2011). From SOMAmer-Based Biomarker Discovery to Diagnostic and Clinical Applications: A SOMAmer-Based, Streamlined Multiplex Proteomic Assay. *PLOS ONE*, **6**, e26332.

Lambert, J.C., Ibrahim-Verbaas, C.A., Harold, D., Naj, A.C., Sims, R., Bellenguez, C., Jun, G., DeStefano, A.L., Bis, J.C., Beecham, G.W., Grenier-Boley, B., Russo, G., Thornton-Wells, T.A., Jones, N., Smith, A.V., Chouraki, V., Thomas, C., Ikram, M.A., Zelenika, D., Vardarajan, B.N., Kamatani, Y., Lin, C.F., Gerrish, A., Schmidt, H., Kunkle, B., Dunstan, M.L., Ruiz, A., Bihoreau, M.T., Choi, S.H., Reitz, C., Pasquier, F., Hollingworth, P., Ramirez, A., Hanon, O., Fitzpatrick, A.L., Buxbaum, J.D., Campion, D., Crane, P.K., Baldwin, C., Becker, T., Gudnason, V., Cruchaga, C., Craig, D., Amin, N., Berr, C., Lopez, O.L., De Jager, P.L., Deramecourt, V., Johnston, J.A., Evans, D., Lovestone, S., Letenneur, L., Morón, F.J., Rubinsztein, D.C., Eiriksdottir, G., Sleegers, K., Goate, A.M., Fiévet, N., Huentelman, M.J., Gill, M., Brown, K., Kamboh, M.I., Keller, L., Barberger-Gateau, P., McGuinness, B., Larson, E.B., Green, R., Myers, A.J., Dufouil, C., Todd, S., Wallon, D., Love, S., Rogaeva, E., Gallacher, J., George-Hyslop, P.S., Clarimon, J., Lleo, A., Bayer, A., Tsuang, D.W., Yu, L., Tsolaki, M., Bossù, P., Spalletta, G., Proitsi, P., Collinge, J., Sorbi, S., Sanchez-Garcia, F., Fox, N.C., Hardy, J., Naranjo, M.C.D., Bosco, P., Clarke, R., Brayne, C., Galimberti, D., Mancuso, M., Matthews, F., Moebus, S., Mecocci, P., Del Zompo, M., Maier, W., Hampel, H., Pilotto, A., Bullido, M., Panza, F., Caffarra, P., Nacmias, B., Gilbert, J.R., Mayhaus, M., Lannfelt, L., Hakonarson, H., Pichler, S., Carrasquillo, M.M., Ingelsson, M., Beekly, D., Alvarez, V., Zou, F., Valladares, O., Younkin, S.G., Coto, E., Hamilton-Nelson, K.L., Gu, W., Razquin, C., Pastor, P., Mateo, I., Owen, M.J., Faber, K.M., Jonsson, P.V., Combarros, O., O'Donovan, M.C., Cantwell, L.B., Soininen, H., Blacker, D., Mead, S., Mosley, T.H., Bennett, D.A., Harris, T.B., Fratiglioni, L., Holmes, C., de Bruijn, R.F.A.G., Passmore, P., Montine, T.J., Bettens, K., Rotter, J.I., Brice, A., Morgan, K., Foroud, T.M., Kukull, W.A., Hannequin, D., Powell, J.F., Nalls, M.A., Ritchie, K., Lunetta, K.L., Kauwe, J.S.K., Boerwinkle, E., Riemenschneider, M., Boada, M., Hiltunen, M., Martin, E.R., Schmidt, R., Rujescu, D., Wang, L.S., Dartigues, J.F., Mayeux, R., Tzourio, C., Hofman, A., Nöthen, M.M., Graff, C., Psaty, B.M., Jones, L., Haines, J.L., Holmans, P.A., Lathrop, M., Pericak-Vance, M.A., Launer, L.J.,

Farrer, L.A., van Duijn, C.M., Van Broeckhoven, C., Moskvina, V., Seshadri, S., Williams, J., Schellenberg, G.D. & Amouyel, P. (2013). Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nature Genetics*, **45**, 1452–1458.

Landau, S.M., Lu, M., Joshi, A.D., Pontecorvo, M.J., Mintun, M.A., Trojanowski, J.Q., Shaw, L.M., Jagust, W.J. & for ADNI (2013). Comparing PET imaging and CSF measurements of Aβ. *Annals of Neurology*, **74**, 826–836.

Lee, S.H., Harold, D., Nyholt, D.R., Consortium, A., Consortium, I.E., disease Consortium, G., for Alzheimer's, E.R., Goddard, M.E., Zondervan, K.T., Williams, J., Montgomery, G.W., Wray, N.R. & Visscher, P.M. (2013). Estimation and partitioning of polygenic variation captured by common SNPs for Alzheimer's disease, multiple sclerosis and endometriosis. *Human Molecular Genetics*, **22**, 832–841.

Ling, H., Hardy, J. & Zetterberg, H. (2015). Neurological consequences of traumatic brain injuries in sports. *Molecular and Cellular Neuroscience*, **66**, 114–122.

Liu, C.C., Liu, C.C., Kanekiyo, T., Xu, H. & Bu, G. (2013). Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. *Nature Reviews Neurology*, **9**, 106–18.

Liu, E., Schmidt, M.E., Margolin, R., Sperling, R., Koeppe, R., Mason, N.S., Klunk, W.E., Mathis, C.A., Salloway, S., Fox, N.C., Hill, D.L., Les, A.S., Collins, P., Gregg, K.M., Di, J., Lu, Y., Tudor, I.C., Wyman, B.T., Booth, K., Broome, S., Yuen, E., Grundman, M. & Brashear, H.R. (2015). Amyloid-β 11C-PiB-PET imaging results from 2 randomized bapineuzumab phase 3 AD trials. *Neurology*, **85**, 692–700.

Louveau, A., Smirnov, I., Keyes, T.J., Eccles, J.D., Rouhani, S.J., Peske, J.D., Derecki, N.C., Castle, D., Mandell, J.W., Lee, K.S., Harris, T.H. & Kipnis, J. (2015). Structural and functional features of central nervous system lymphatic vessels. *Nature*, **523**, 337–341.

Lunnon, K., Ibrahim, Z., Proitsi, P. & Lourdusamy, A. (2012). Mitochondrial Dysfunction and Immune Activation are Detectable in Early Alzheimer's Disease Blood. *Journal of Alzheimer's Disease*, **30**, 685–710.

Mapstone, M., Cheema, A.K., Fiandaca, M.S., Zhong, X., Mhyre, T.R., MacArthur, L.H., Hall, W.J., Fisher, S.G., Peterson, D.R., Haley, J.M., Nazar, M.D., Rich, S.A., Berlau, D.J., Peltz, C.B., Tan, M.T., Kawas, C.H. & Federoff, H.J. (2014). Plasma phospholipids identify antecedent memory impairment in older adults. *Nature Medicine*, **20**, 415–418.

Marques, F., Sousa, J.o.C., Sousa, N. & Palha, J.A. (2013). Blood-brain-barriers in aging and in Alzheimer's disease. *Molecular Neurodegeneration*, **8**, 38.

Matthews, F.E., Stephan, B.C.M., Robinson, L., Jagger, C., Barnes, L.E., Arthur, A., Brayne, C., Comas-Herrera, A., Wittenberg, R., Dening, T., McCracken, C., Moody, C., Parry, B., Green, E., Barnes, R., Warwick, J., Gao, L., Mattison, A., Baldwin, C., Harrison, S., Woods, B., McKeith, I., Ince, P., Wharton, S. & Forster, G. (2016). A two decade dementia incidence comparison from the Cognitive Function and Ageing Studies I and II. *Nature Communications*, **7**, 11398.

Mattsson, N., Zetterberg, H. & Blennow, K. (2010). Lessons from Multicenter Studies on CSF Biomarkers for Alzheimer's Disease. *International journal of Alzheimer's disease*, **2010**, 1–5.

Maurer, K., Volk, S. & Gerbaldo, H. (1997). Auguste D and Alzheimer's disease. *Lancet*, **349**, 1546–1549.

Menéndez-González, M. (2014). Routine lumbar puncture for the early diagnosis of Alzheimer's disease. Is it safe? *Frontiers in Aging Neuroscience*, **6**, 1–2.

Montagne, A., Barnes, S.R., Sweeney, M.D., Halliday, M.R., Sagare, A.P., Zhao, Z., Toga, A.W., Jacobs, R.E., Liu, C.Y., Amezcua, L., Harrington, M.G., Chui, H.C., Law, M. & Zlokovic, B.V. (2015). Blood-Brain Barrier Breakdown in the Aging Human Hippocampus. *Neuron*, **85**, 296–302.

Morris, G.P., Clark, I.a., Zinn, R. & Vissel, B. (2013). Microglia: A new frontier for synaptic plasticity, learning and memory, and neurodegenerative disease research. *Neurobiology of Learning and Memory*, **105**, 40–53.

Morris, G.P., Clark, I.a. & Vissel, B. (2014). Inconsistencies and controversies surrounding the amyloid hypothesis of Alzheimer's disease. *Acta neuropathologica communications*, **2**, 135.

Nho, K., Corneveaux, J.J., Kim, S., Lin, H., Risacher, S.L., Shen, L., Swaminathan, S., Ramanan, V.K., Liu, Y., Foroud, T., Inlow, M.H., Siniard, A.L., Reiman, R.A., Aisen, P.S., Petersen, R.C., Green, R.C., Jack, C.R., Weiner, M.W., Baldwin, C.T., Lunetta, K., Farrer, L.A., Furney, S.J., Lovestone, S., Simmons, A., Mecocci, P., Vellas, B., Tsolaki, M., Kloszewska, I., Soininen, H., McDonald, B.C., Farlow, M.R., Ghetti, B., Huentelman, M.J. & Saykin, A.J. (2013). Whole-exome sequencing and imaging genetics identify functional variants for rate of change in hippocampal volume in mild cognitive impairment. *Molecular psychiatry*, **18**, 781–787.

Nho, K., Kim, S., Risacher, S.L., Shen, L., Corneveaux, J.J., Swaminathan, S., Lin, H., Ramanan, V.K., Liu, Y., Foroud, T.M., Inlow, M.H., Siniard, A.L., Reiman, R.a., Aisen, P.S., Petersen, R.C., Green, R.C., Jack, C.R., Weiner, M.W., Baldwin, C.T., Lunetta, K.L., Farrer, L.a., Furney, S.J., Lovestone, S., Simmons, A., Mecocci, P., Vellas, B., Tsolaki, M., Kloszewska, I., Soininen, H., McDonald, B.C., Farlow, M.R., Ghetti, B., Huentelman, M.J. & Saykin, A.J. (2015). Protective variant for hippocampal atrophy identified by whole exome sequencing. *Annals of Neurology*, **77**, 547–552.

Norton, S., Matthews, F.E., Barnes, D.E., Yaffe, K. & Brayne, C. (2014). Potential for primary prevention of Alzheimer's disease: an analysis of population-based data. *The Lancet Neurology*, **13**, 788–794.

O'Bryant, S. & Xiao, G. (2014). Validation of a Serum Screen for Alzheimer's Disease Across Assay Platforms, Species, and Tissues. *Journal of Alzheimer's disease*, **42**, 1325–1335.

O'Bryant, S.E., Gupta, V., Henriksen, K., Edwards, M., Jeromin, A., Lista, S., Bazenet, C., Soares, H., Lovestone, S., Hampel, H., Montine, T., Blennow, K., Foroud, T., Carrillo, M., Graff-Radford, N., Laske, C., Breteler, M., Shaw, L., Trojanowski, J.Q., Schupf, N., Rissman, R.A., Fagan, A.M., Oberoi, P., Umek, R., Weiner, M.W., Grammas, P., Posner, H. & Martins, R. (2015). Guidelines for the standardization of preanalytic variables for blood-based biomarker studies in Alzheimer's disease research. *Alzheimer's & dementia : the journal of the Alzheimer's Association*, **11**, 549–560.

Okello, A., Koivunen, J., Edison, P., Archer, H.A., Turkheimer, F.E., N??gren, K., Bullock, R., Walker, Z., Kennedy, A., Fox, N.C., Rossor, M.N., Rinne, J.O. & Brooks, D.J. (2009). Conversion of amyloid positive and negative MCI to AD over 3 years: an 11C-PIB PET study. *Neurology*, **73**, 754–760.

Pedersen, N.L., Gatz, M., Berg, S. & Johansson, B. (2004). How Heritable Is Alzheimer's Disease Late in Life? Findings from Swedish Twins. *Annals of Neurology*, **55**, 180–185.

Proitsi, P., Kim, M., Whiley, L., Pritchard, M., Leung, R., Soininen, H., Kloszewska, I., Mecocci, P., Tsolaki, M., Vellas, B., Sham, P., Lovestone, S., Powell, J.F., Dobson, R.J.B. & Legido-Quigley, C. (2015). Plasma lipidomics analysis finds long chain cholesteryl esters to be associated with Alzheimer's disease. *Translational Psychiatry*, **5**, e494–e498.

Purcell, S.M., Wray, N.R., Stone, J.L., Visscher, P.M., O'Donovan, M.C., Sullivan, P.F., Sklar, P., Stone, J.L., Ruderfer, D.M., McQuillin, A., Morris, D.W., O'Dushlaine, C.T., Corvin, A., Holmans, P.A., Sklar, P., Macgregor, S., Sklar, P., Gurling, H., Blackwood, D.H.R., Corvin, A., Craddock, N.J., Gill, M., Hultman, C.M., Kirov, G.K., Lichtenstein, P., Muir, W.J., Owen, M.J., Pato, C.N., Scolnick, E.M., St Clair, D., Sklar Leader, P., Williams, N.M., Georgieva, L., Nikolov, I., Norton, N., Williams, H., Toncheva, D., Milanova, V., Thelander, E.F., Sullivan, P., Kenny, E., Quinn, E.M., Choudhury, K., Datta, S., Pimm, J., Thirumalai, S., Puri, V., Krasucki, R., Lawrence, J., Quested, D., Bass, N., Crombie, C., Fraser,

G., Leh Kuan, S., Walker, N., McGhee, K.A., Pickard, B., Malloy, P., Maclean, A.W., Van Beck, M., Pato, M.T., Medeiros, H., Middleton, F., Carvalho, C., Morley, C., Fanous, A., Conti, D., Knowles, J.A., Paz Ferreira, C., Macedo, A., Helena Azevedo, M., Kirby, A.N., Ferreira, M.A.R., Daly, M.J., Chambert, K., Kuruvilla, F., Gabriel, S.B., Ardlie, K. & Moran, J.L. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, **460**, 748–752.

Qiu, C., Kivipelto, M. & Von Strauss, E. (2009). Epidemiology of Alzheimer's disease: Occurrence, determinants, and strategies toward intervention. *Dialogues in Clinical Neuroscience*, **11**, 111–128.

Räihä, I., Kaprio, J., Koskenvuo, M., Rajala, T. & Sourander, L. (1998). Environmental differences in twin pairs discordant for Alzheimer's disease. *Journal of neurology, neurosurgery, and psychiatry*, **65**, 785–7.

Ramirez-Alvarado, M., Merkel, J.S. & Regan, L. (2000). A systematic exploration of the influence of the protein stability on amyloid fibril formation in vitro. *PNAS*, **97**, 8979–8984.

Ridge, P.G., Mukherjee, S., Crane, P.K., Kauwe, J.S.K. & Consortium, A.D.G. (2013). Alzheimer's Disease: Analyzing the Missing Heritability. *PLOS ONE*, **8**, e79771.

Ridge, P.G., Hoyt, K.B., Boehme, K., Mukherjee, S., Crane, P.K., Haines, J.L., Mayeux, R., Farrer, L.A., Pericak-Vance, M.A., Schellenberg, G.D. & Kauwe, J.S.K. (2016). Assessment of the genetic variance of late-onset Alzheimer's disease. *Neurobiology of aging*, **41**, 200.e13–200.e20.

Rockwood, K. (2004). Size of the treatment effect on cognition of cholinesterase inhibition in Alzheimer's disease. *Journal of neurology, neurosurgery, and psychiatry*, **75**, 677–685.

Rosen, W.G., Mohs, R.C. & Davis, K.L. (1984). A new rating scale for Alzheimer's disease. *American Journal of Psychiatry*, **141**, 1356–1364,.

Sabuncu, M.R., Buckner, R.L., Smoller, J.W., Lee, P.H., Fischl, B., Sperling, R.A. & for ADNI (2012). The Association between a Polygenic Alzheimer Score and Cortical Thickness in Clinically Normal Subjects. *Cerebral Cortex*, **22**, 2653–2661.

Sachdev, P.S., Blacker, D., Blazer, D.G., Ganguli, M., Jeste, D.V., Paulsen, J.S. & Petersen, R.C. (2014). Classifying neurocognitive disorders: the DSM-5 approach. *Nature reviews Neurology*, **10**, 634–642.

Salloway, S., Sperling, R., Fox, N.C., Blennow, K., Klunk, W., Raskind, M., Sabbagh, M., Honig, L.S., Porsteinsson, A.P., Ferris, S., Reichert, M., Ketter, N., Nejadnik, B., Guenzler, V., Miloslavsky, M., Wang, D., Lu, Y., Lull, J., Tudor, I.C., Liu, E., Grundman, M., Yuen, E., Black, R. & Brashear, H.R. (2014). Two Phase 3 Trials of Bapineuzumab in Mild-to-Moderate Alzheimer's Disease. *New England Journal of Medicine*, **370**, 322–333.

Salmon, D.P. & Bondi, M.W. (2010). Neuropsychological Assessment of Dementia. *Annual Review of Psychology*, **60**, 257–282.

Scheinin, N.M., Aalto, S., Kaprio, J., Koskenvuo, M., Räihä, I., Rokka, J., Hinkka-Yli-Salomäki, S. & Rinne, J.O. (2011). Early detection of Alzheimer disease. *Neurologyy*, **77**, 453–460.

Sheridan, C. (2014). Milestone approval lifts Illumina's NGS from research into clinic. *Nature biotechnology*, **32**, 111–112.

Sims, R., van der Lee, S.J., Naj, A.C., Vronskaya, M. & Badarinarayan, N. (2015). Characterizing missing heritability of late-onset Alzheimer's disease: An exome array study. *Alzheimer's & Dementia*, **11**, P856–P857.

Sleegers, K., Bettens, K., De Roeck, A., Van Cauwenberghe, C., Cuyvers, E., Verheijen, J., Struyfs, H., Van Dongen, J., Vermeulen, S., Engelborghs, S., Vandenbulcke, M., Vandenberghe, R., De Deyn, P.P., Van Broeckhoven, C. & Consortium, B. (2015). A 22-single nucleotide polymorphism Alzheimer risk score correlates with family history, onset age, and cerebrospinal fluid Aβ42. *Alzheimer's & Dementia*, **11**, 1452–1460.

Snyder, H.M., Carrillo, M.C., Grodstein, F., Henriksen, K., Jeromin, A., Lovestone, S., Mielke, M.M., O'Bryant, S., Sarasa, M., Sjögren, M., Soares, H., Teeling, J., Trushina, E., Ward, M., West, T., Bain, L.J., Shineman, D.W., Weiner, M. & Fillit, H.M. (2014). Developing novel blood-based biomarkers for Alzheimer's disease. *Alzheimer's & Dementia*, **10**, 109–114.

Song, F., Parekh, S., Hooper, L., Loke, Y.K., Ryder, J., Sutton, A.J., Hing, C., Kwok, C.S., Pang, C. & Harvey, I. (2010). Dissemination and publication of research findings: An updated review of related biases. *Health Technology Assessment*, **14**, 1–220.

Sperling, R.A., Rentz, D.M., Johnson, K.A., Karlawish, J., Donohue, M., Salmon, D.P. & Aisen, P. (2014). The A4 Study: Stopping AD Before Symptoms Begin? *Science Translational Medicine*, **6**, 228.

Tao, J., Chiam, W., Lunnon, K., Voyle, N., Proitsi, P. & Coppola, G. (2015). No Evidence to Suggest that the Use of Acetylcholinesterase Inhibitors Confounds the Results of Two Blood-Based Biomarker Studies in Alzheimer's Disease. *Journal of Alzheimer's Disease*, **47**, 741–750.

Thal, D.R., Attems, J. & Ewers, M. (2014). Spreading of amyloid, tau, and microvascular pathology in alzheimer's disease: Findings from neuropathological and neuroimaging studies. *Journal of Alzheimer's Disease*, **42**, S421–S429.

Thambisetty, M., An, Y., Kinsey, A., Koka, D., Saleem, M., Gntert, A., Kraut, M., Ferrucci, L., Davatzikos, C., Lovestone, S. & Resnick, S.M. (2012). Plasma clusterin concentration is associated with longitudinal brain atrophy in mild cognitive impairment. *NeuroImage*, **59**, 212–217.

Touw, W.G., Bayjanov, J.R., Overmars, L., Backus, L., Boekhorst, J., Wels, M. & van Hijum, S.A. (2012). Data mining in the Life Sciences with Random Forest : a walk in the park or lost in the jungle ? *Briefings in Bioinformatics*, **14**, 315–326.

Villemagne, V.L., Pike, K.E., Chételat, G., Ellis, K.A., Mulligan, R.S., Bourgeat, P., Ackermann, U., Jones, G., Szoeke, C., Salvado, O., Martins, R., O'Keefe,

G., Mathis, C.A., Klunk, W.E., Ames, D., Masters, C.L. & Rowe, C.C. (2011). Longitudinal assessment of Aβ and cognition in aging and Alzheimer disease. *Annals of neurology*, **69**, 181–192.

Viña, J. & Lloret, A. (2010). Why women have more Alzheimer's disease than men: gender and mitochondrial toxicity of amyloid-beta peptide. *Journal of Alzheimer's disease*, **20**, S527–33.

Voyle, N., Baker, D., Burnham, S.C., Covin, A., Zhang, Z.P., Sangurdekar, D.P., Hehir, C.A.T., Bazenet, C., Lovestone, S., Kiddle, S. & Dobson, R.J.B. (2015). Blood Protein Markers of Neocortical Amyloid-beta Burden: A Candidate Study Using SOMAscan Technology. *Journal of Alzheimers Disease*, **46**, 947–961.

Voyle, N., Keohane, A., Newhouse, S., Lunnon, K., Johnston, C., Soininen, H., Kloszewska, I., Mecocci, P., Tsolaki, M., Vellas, B., Lovestone, S., Hodges, A., Kiddle, S. & Dobson, R.J.B. (2016a). A Pathway Based Classification Method for Analyzing Gene Expression for Alzheimer' s Disease Diagnosis. *Journal of Alzheimer's Disease*, **49**, 659–669.

Voyle, N., Kim, M., Proitsi, P., Ashton, N.J., Baird, A.L., Bazenet, C., Hye, A., Westwood, S., Chung, R., Ward, M., Rabinovici, G.D., Lovestone, S., Breen, G., Legido-Quigley, C., Dobson, R.J.B. & Kiddle, S.J. (2016b). Blood metabolite markers of neocortical amyloid-β burden: discovery and enrichment using candidate proteins. *Translational Psychiatry*, **6**, e719.

Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., Klemm, A., Flicek, P., Manolio, T., Hindorff, L. & Parkinson, H. (2014). The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Research*, **42**, 1001–1006.

Westman, E., Simmons, A., Muehlboeck, J.S., Mecocci, P., Vellas, B., Tsolaki, M., Kłoszewska, I., Soininen, H., Weiner, M.W., Lovestone, S., Spenger, C., Wahlund, L.O. & for the AddNeuroMed consortium and ADNI (2011). AddNeuroMed and

ADNI: Similar patterns of Alzheimer's atrophy and automated MRI classification accuracy in Europe and North America. *NeuroImage*, **58**, 818–828.

Whiley, L., Sen, A., Heaton, J., Proitsi, P., García-Gómez, D., Leung, R., Smith, N., Thambisetty, M., Kloszewska, I., Mecocci, P., Soininen, H., Tsolaki, M., Vellas, B., Lovestone, S., Legido-Quigley, C. & Consortium, A. (2014). Evidence of altered phosphatidylcholine metabolism in Alzheimer's disease. *Neurobiology of Aging*, **35**, 271–278.

Whiteley, W., Tian, Y. & Jickling, G.C. (2012). Blood biomarkers in stroke : research and clinical practice. *International Journal of Stroke*, **7**, 435–439.

Wimo, A., Jönsson, L., Bond, J., Prince, M., Winblad, B. & International, A.D. (2013). The worldwide economic impact of dementia 2010. *Alzheimer's & Dementia*, **9**, 1–11.

Wiseman, F.K. (2015). A genetic cause of Alzheimer disease : mechanistic insights from Down syndrome. *Nature reviews. Neuroscience*, **16**, 564–574.

Wold, H. (2004). Partial Least Squares. In *Encyclopedia of Statistical Sciences*.

Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R., Madden, P.A.F., Heath, A.C., Martin, N.G., Montgomery, G.W., Goddard, M.E. & Visscher, P.M. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, **42**, 565–569.

Zhao, Z., Nelson, A.R., Betsholtz, C. & Zlokovic, B.V. (2015). Establishment and Dysfunction of the Blood-Brain Barrier. *Cell*, **163**, 1064–1078.

Zürbig, P. & Jahn, H. (2012). Use of proteomic methods in the analysis of human body fluids in Alzheimer research. *Electrophoresis*, **33**, 3617–3630.

Zwan, M.D., Rinne, J.O., Hasselbalch, S.G., Nordberg, A., Lleo, A., Herukka, S.K., Soininen, H., Law, I., Bahl, J.M.C., Carter, S.F., Fortea, J., Blesa, R., Teunissen, C.E., Bouwman, F.H., van Berckel, B.N.M. & Visser, P.J. (2015). Use of amyloid-

PET to determine cutpoints for CSF markers: A multicenter study. *American Academy of Neurology*, **86**, 50–58.

Appendices

Appendix A

No Evidence to Suggest that the Use of Acetylcholinesterase Inhibitors Confounds the Results of Two Blood-Based Biomarker Studies in Alzheimers Disease

Journal of Alzheimer's Disease 47 (2015) 741–750
DOI 10.3233/JAD-150289
IOS Press

No Evidence to Suggest that the Use of Acetylcholinesterase Inhibitors Confounds the Results of Two Blood-Based Biomarker Studies in Alzheimer's Disease

Justin Tao Wen Chiam[a], Katie Lunnon[b], Nicola Voyle[c,e], Petroula Proitsi[c,d], Giovanni Coppola[f], Daniel Geschwind[f], Sally Nelson[g], Caroline Johnston[c,d], Hilkka Soininen[h], Iwona Kłoszewska[i], Patrizia Mecocci[j], Magda Tsolaki[k], Bruno Vellas[l], Angela Hodges[c,d], Simon Lovestone[c,m], Stephen Newhouse[c,d], Richard James Butler Dobson[c,d], Steven John Kiddle[c,e,1] and Martina Sattlecker[c,d,1,*]

[a]*Guy's, King's and St Thomas School of Medical Education, King's College London, London, UK*
[b]*Institute of Clinical and Biomedical Science, University of Exeter, Exeter, Devon, UK*
[c]*Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK*
[d]*NIHR Biomedical Research Centre for Mental Health and Biomedical Research, London, UK*
[e]*MRC Social, Genetic and Developmental Psychiatry Centre, King's College London, London, UK*
[f]*Department of Neurology, Programme in Neurogenetics, David Geffen School of Medicine, University of California at Los Angeles, CA, USA*
[g]*SomaLogic, Boulder, CO, USA*
[h]*Department of Neurology, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland*
[i]*Medical University of Lodz, Lodz, Poland*
[j]*Institute of Gerontology and Geriatrics, University of Perugia, Perugia, Italy*
[k]*3rd Department of Neurology, "G. Papanicolaou" Hospital, Aristotle University of Thessaloniki, Thessaloniki, Greece*
[l]*INSERM U 558, University of Toulouse, Toulouse, France*
[m]*Department of Psychiatry, University of Oxford, Oxford, UK*

Accepted 9 May 2015

Abstract.
Background: There is an urgent need to discover Alzheimer's disease (AD) biomarkers that are both easily measured and reliable. Research into blood-based biomarkers for AD using transcriptomics and proteomics has been an attractive and promising area of research. However, to date researchers have not looked into the possibility of AD medication being a confounding factor in these studies.
Objective: This study explored whether acetylcholinesterase inhibitors (AChEIs), the main class of AD medication, are a confounding factor in AD blood biomarker studies.
Methods: The most promising blood transcriptomic and proteomic biomarkers from two recent studies were analyzed to determine if they were differentially expressed between AD subjects on AChEIs and subjects that were not.

[1]These authors contributed equally to this work.

*Correspondence to: Martina Sattlecker, NIHR Biomedical Research Centre for Mental Health South London and Maudsley NHS Foundation Trust & King's College London, Institute of Psychiatry, De Crespigny Park, London, UK. Tel.: +44 20 7848 0236; E-mail: martina.sattlecker@kcl.ac.uk.

ISSN 1387-2877/15/$35.00

Results: None of the gene or protein biomarkers analyzed were found to be significantly altered between subjects in either group.

Conclusion: This study found no evidence that AChEIs are a confounding factor in these published AD blood biomarker studies. Further work is needed to confirm that this is also the case for other proposed biomarkers.

Keywords: Alzheimer's disease, blood, cholinesterase inhibitors, gene expression, microarray, protein, proteomics

INTRODUCTION

Alzheimer's disease (AD) is a common, costly, and fatal neurodegenerative disorder. It manifests in the form of progressive cognitive decline, including memory loss, executive dysfunction, psychiatric symptoms, and behavioral disturbances [1]. Currently, a definitive diagnosis of AD can only be obtained after postmortem dissection of brain tissue. AD diagnosis therefore relies on robust clinical evaluation, sometimes including assessment of pathology using cerebrospinal fluid (CSF) biomarkers [2], or pathology measures from brain scans; usually magnetic resonance imaging (MRI) or positron emission tomography (PET) [3, 4].

Despite these efforts, diagnosis of AD remains difficult, especially in areas where access to advanced neuroimaging equipment remains limited. A lumbar puncture to acquire CSF for analysis is also relatively invasive and inappropriate in certain clinical environments. It has been suggested that up to two-thirds of dementia patients go undiagnosed [5], and that by the time an AD diagnosis is made, the underlying pathological processes have been developing for around 20 years [6]. There is therefore an urgent need to develop investigative techniques that are cost effective, easy to administer, and capable of aiding the diagnosis of AD in its early stages. Though undoubtedly challenging, this may prove useful for enriching clinical trials for subjects whose pathology is less advanced.

Recently, the analysis of blood samples to develop a blood-based diagnostic test has been an attractive area in AD biomarker research. Blood samples of AD subjects can be collected with relative ease, and analyzed to determine differences in protein or messenger RNA (mRNA) quantity that might elucidate underlying biological changes in the disease state [5]. Studies such as Booij et al. [7], Fehlbaum-Beurdeley et al. [8], and more recently Lunnon et al. [9] have demonstrated that whole-blood profiling of mRNA can generate evidence of AD associated differences in gene expression. Similarly, studies such as Ray et al. [10], Doecke et al. [11], and Sattlecker et al. [12] have analyzed blood protein quantities to identify proteins with significantly altered blood levels in AD. Zurbig and Jahn [13], Lista et al. [14], Kiddle et al. [15], and Chiam et al. [16] have also recently reviewed the blood-based proteins most commonly associated with AD, finding a modest degree of replication between studies. Given the prospect of a cheap and convenient diagnostic test should success be achieved, it is likely that research into blood-based biomarkers, while not without limitations, will continue to be an appealing avenue in AD biomarker discovery [17, 18].

Although there has been a significant amount of research generated in the field, very little has been done to look into the potential effects of medication as a confounding factor in blood-based biomarker discovery. This is important as the medication an individual is receiving depends largely on their diagnosis. Medication could potentially affect the composition of biological molecules in the blood. For example, AD patients are often placed on a wide range of psychotropic drugs and recent evidence suggests that psychotropic drugs affect the expression of AD related genes in blood [19].

There are currently two classes of drugs used to treat the cognitive symptoms of AD, acetylcholinesterase inhibitors (AChEIs) and memantine (an NMDA receptor antagonist). Of these, AChEIs are the most common class of drug used in the treatment of AD [20]. These drugs act by inhibiting the enzyme acetylcholinesterase. This prevents the breakdown of the neurotransmitter acetylcholine at synapses, thus increasing the strength of neural transmission in the brain [21, 22]. Treatment with AChEIs has been found to modulate the expression of pro- and anti-inflammatory cytokines in the blood of AD patients [23]. Given the high prevalence of AChEI use among AD subjects, it is therefore pertinent to determine if the potential gene and protein blood biomarkers identified in studies are indeed due to biological changes as a result of AD and not due to the effects of these drugs.

This study seeks to build on the results from Lunnon et al. [9] and Sattlecker et al. [12] by analyzing the most promising gene and protein biomarkers identified in AD subjects by both studies and determining if the use of AChEIs confounds the association between these blood-based biomarkers and AD.

METHODS AND MATERIALS

Samples and clinical data from the AddNeuroMed cohort

As described in Lunnon et al. [9] and Sattlecker et al. [12], blood samples were obtained from subjects participating in the AddNeuroMed (ANM) study [24–26]. Subjects were located at six different study sites across Europe, namely London, Lodz, Toulouse, Perugia, Kuopio, and Thessaloniki. Informed consent was appropriately taken according to the Declaration of Helinski (1991) and ethical approval was obtained at each site. A diagnosis of AD was attained using the NINCDS-ADRDA criteria [27] and the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV).

All subjects went through a semi-structured interview in order to collect the necessary demographic and medical information. This included an array of neuropsychological assessments such as the Mini-Mental State Examination [28], Global Deterioration Scale [29] and the Alzheimer's Disease Assessment Scale-Cognitive subscale [30]. Information regarding the kind of neuropsychiatric medication each subject was currently prescribed was also obtained.

Blood samples were drawn from subjects by venipuncture and collected in PAXgen vacutainer tubes (Qiagen) for RNA analysis and EDTA glass tubes for proteomic analysis.

Relevant demographics and clinical data for these subjects were extracted from the AddNeuroMed database using CohortExplorer [31].

Gene expression study

Blood samples collected from the ANM cohort were analyzed to determine mRNA gene expression profiles. Full details of the data collection and preprocessing procedure can be found in Lunnon et al. [9] and Voyle et al. (manuscript in preparation), and is only discussed briefly here.

Data collection

The vacutainer tubes containing blood samples for RNA analysis were inverted 8–10 times and stored at –24°C for 24 h before lowering the temperature to –80°C until RNA extraction. RNA was extracted using the PAXgene blood RNA kit (Qiagen), following the manufacturer's protocol. The 2100 Bioanalyser (Agilent Technologies) was then used to evaluate the quality of the extracted RNA. Only RNA samples that exceeded a RNA integrity number (RIN) of 7.0 were used in the analysis.

Microarray processing was conducted at the University of California in Los Angeles. The RNA samples were processed on Illumina Human HT-12 v3 Expression BeadChips (Illumina), each containing 48,803 probes. RNA was amplified using the Illumina TotalPrep RNA Amplification Kit (Ambion) and gene expression values were obtained using the Lumi package within the R Bioconductor project [32].

Preprocessing

Raw gene expression data was subject to a model based background correction for bead array [33]. Negative bead expression levels were used to correct for background noise. The data was then log base 2 transformed and robust spline normalized before outlying samples were iteratively identified by fundamental network concepts and removed [32, 34]. To remove any batch effects, we adjusted for technical categorical variables using ComBat [35]. The first principal component across housekeeping probes was taken and regressed against technical variables and phenotype in order to account for the principal component across housekeeping probes. Variables significantly associated with the first principal component where then regressed against expression for each probe, and the mean adjusted residuals taken forward for all further analyses. Finally, the data was subset to probes that could be reliably detected in at least 80% of samples in each diagnostic group. Subjects were excluded if lab-based investigations highlighted discrepancies between recorded sex and sex determined by the XIST gene.

Probe selection

Probes that had a significant difference in signal between AD subjects and controls were identified from Lunnon et al. [9]. The most significant probes were identified through a dual-criteria. Firstly a Bonferroni correction was applied to the *p*-values reported by Lunnon et al. [9] ($\alpha = 0.05/19{,}161$). Any probe that met the corrected threshold of *p*-value $<\alpha$ and had an absolute fold change of >0.5 was identified for further analysis.

Subject selection

In Lunnon et al. [9], a total of 356 samples (116 control, 127 mild cognitive impairment (MCI), 113 AD) were put through RNA microarray processing. Of this a subset was extracted for analysis in our study, as described below. Of the 113 AD subjects, 100 had complete demographic data for medication, age, gender,

Mini-Mental State Examination (MMSE) score, and *APOE* status. It was decided that subjects on memantine, another form of AD cognitive drug would be excluded as there were too few subjects to conduct a conclusive analysis. Therefore 5 subjects on memantine and 6 subjects on both AChEI and memantine were excluded. This resulted in a total of 89 AD subjects for gene expression analysis.

Proteomic study

Similar to the gene expression analysis described above, blood samples collected from the ANM cohort were analyzed to determine protein quantities. A detailed account of the data collection and preprocessing procedure can be found in Sattlecker et al. [12] and is described only briefly here.

Data collection

EDTA tubes with the blood samples were centrifuged at 2,000 rpm at 4°C for 10 min within approximately 2 h of collection. The resulting plasma supernatant was then collected and divided into aliquots before being frozen at –80°C until protein measurement.

Protein quantities were measured using the new Slow Off-rate Modified Aptamer (SOMAmer)-based capture array known as "SOMAscan" (SomaLogic, Inc). Chemically modified nucleotides are used to transform a protein signal into a nucleotide signal [36]. Microarrays were then used to quantify the signal using relative fluorescence. A total of 1,001 human proteins, representing different molecular pathways and gene families, were quantified this way.

Preprocessing

Hybridization controls on the microarray were used to monitor sample-by-sample variability in hybridization, while the median signal across all SOMAmers was used to monitor overall technical variability. Using both the resulting hybridization and median scale factors, data across samples was normalized. An acceptance criterion of 0.4–2.5 was used for values based on historic trends. SOMAmer-by-SOMAmer calibration was established through the repeated measurement of calibration samples. A calibration scale factor is then generated using historic values of these calibration samples. The acceptance criterion for calibrator scale factors is that 95% of SOMAmers must have a calibration scale factor within $\pm$ 0.4 of the median [12].

All measurements obtained were log 2 transformed. Seven sample outliers were identified using principal component analysis in R and were thus removed from downstream processing [12].

Probe selection

Sattlecker et al. [12] had previously identified four proteins (prostate-specific antigen complexed to α1-antichymotrypsin, clusterin, pancreatic prohormone and fetuin B) that were found to have significantly altered levels (q-value <0.05) in AD subjects when compared to healthy elderly controls. These four proteins were selected for the analysis.

Subject selection

In Sattlecker et al. [12], a total of 415 ANM samples (110 control, 109 MCI, 196 AD) underwent SOMAscan proteomic analysis. Like in the gene expression portion of the study, a subset of this was extracted for data analysis using the same criteria. Of the 196 AD subjects, 189 had complete demographic data for medication, age, gender, MMSE score, and *APOE* status. After 8 subjects on memantine and 13 subjects on both AChEI and memantine were excluded, a total of 168 AD subjects were selected for the proteomic portion of the study.

Statistical power calculations

Power calculations were performed using the 'pwr.f2.test' in the 'pwr' R package, based on the approach of Cohen [37]. Based on the recommendation by Cohen $f^2 = 0.02$, 0.15, and 0.35 were used to indicate a small, medium, and large effect size, respectively. A statistical significance level of 0.005 was required, to represent a Bonferroni multiple testing correction for 10 markers (midway between number of gene expression and protein markers studied).

Data analysis

All data analysis was performed using R. The AD subjects that were selected for either the gene expression study or the proteomic study were further split into two groups based on the type of cognitive enhancing AD medication they were on. The first group comprised those on AChEIs, while the second group comprised those subjects not on any form of cognitive enhancing AD medication (non-AChEI). The gene expression study had 72 subjects in the AChEI group and 17 subjects in the non-AChEI group. Similarly, the proteomic study had 129 subjects in the AChEI group and 39 subjects in the non-AChEI group. For both studies, an analysis was then conducted between

the AChEI group and non-AChEI group to see if there was any significant difference in gene expression and protein quantity.

Firstly the demographic data was analyzed to ensure that confounding variables could be identified. Discrete variables such as gender and *APOE* ε4 allele status were analyzed using Fisher's exact test. Continuous variables such as age and MMSE score were analyzed using linear modeling. A threshold of p-value <0.05 was set to identify if any of the variables differed between both AChEI and non-AChEI groups.

An analysis was then conducted on the genes and proteins that were pre-selected to determine if there was a significant difference in gene expression or protein levels between AD patients on AChEIs and those that were not. This was conducted using linear modeling with study site added as a covariate in the analysis. The p-values obtained were then adjusted for multiple testing by applying the false discovery rate (FDR). A threshold of q-value <0.05 was used to identify differentially expressed genes and differences in blood protein quantity. Box plots were created for genes or proteins of interest and for these, 111 controls (non-AD subjects) from the ANM cohort were included to provide visual comparison of any variation between the groups.

RESULTS

Statistical power

To examine our ability to detect small, medium, and large effects of AChEIs on blood marker levels, we performed statistical power calculations using a significance level of $p = 0.005$. Figure 1 shows that the gene expression study ($n = 89$) has ∼2.3%, ∼47%, or ∼96% statistical power to detect a small, medium, or large effect of AChEIs on blood gene expression markers. It also shows that the proteomics study ($n = 168$) has ∼5%, ∼89%, or ∼100% statistical power to detect a small, medium, or large effect of AChEIs on blood protein markers. It also shows that even a study with a sample size of 500 would only have ∼30% power to detect a small effect of medication on a blood marker.

Gene expression study

To investigate the effect of AChEIs on blood gene expression markers of AD, we examined gene expression levels in 89 AD subjects with gene expression, demographic, and medication data. No significant differences in gender, age, *APOE* status, or MMSE scores were seen between the AChEI group and the non-AChEI group ($p < 0.05$; Table 1). As a result, none of the demographic variables were factored in as covariates in the analysis of probe signal between both groups.

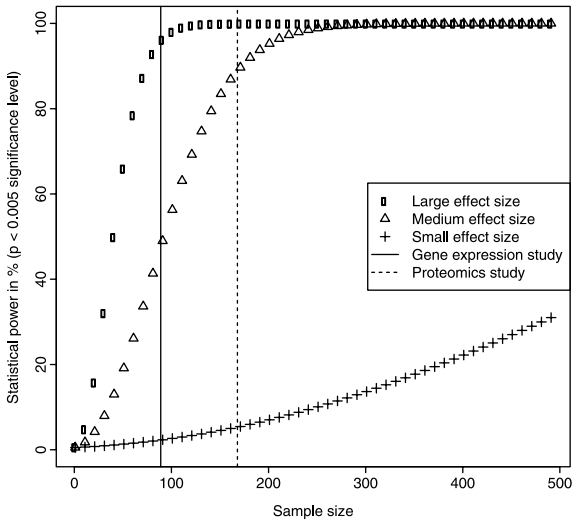


Fig. 1. Statistical power curves for the detection of the effect of AChEIs on 10 blood markers (midway between number of genes and proteins investigated). Points indicate statistical power for different effect sizes (small, medium, and large) in a model with 6 degrees of freedom for the numerator. Black lines indicate sample size of the gene expression study ($n = 89$) and the proteomics study ($n = 168$). A significance threshold of $p < 0.005$ was used to correspond to a $p < 0.05$ significance level after Bonferroni multiple testing correction for 10 markers/tests.

When a Bonferroni correction was applied, 23 probes were identified that passed the dual-criteria that had been set (p-value <α and absolute fold change >0.5). These 23 probes were thus selected for analysis (Table 2).

After multiple testing correction was applied, none of the 23 probes analyzed showed any significant difference in signal strength between the group on AChEIs and the group without (q-value <0.05).

Table 1
Demographic information for subjects included in the gene expression analysis

	AChEI	Non-AChEI	p-value
Total Number	72	17	–
Males/Females	25/47	3/14	0.25
Median Age (IQR)	76 (9.25)	75 (9)	0.89
Median MMSE (IQR)	22 (7)	22 (6)	0.76
APOE ε4 Status (0/1/2)	29/29/14	9/7/1	0.40

AChEI, Group taking acetylcholinesterase inhibitors; Non-AChEI, Group not on any AD cognitive enhancing medication; MMSE, Mini-Mental State Examination.

Table 2
Results for the gene expression analysis showing the 23 probes analyzed with details of the gene name, coefficient value, standard error, *p*-value and q-value

Probe ID	Gene Name	Coefficient	Standard Error	*p*-value	q-value
ILMN_2097421	MRPL51	0.35	0.27	0.20	0.29
ILMN_1784286	NDUFA1	0.40	0.26	0.13	0.21
ILMN_1776104	NDUFS5	0.37	0.27	0.17	0.26
ILMN_1726603	ATP5I	0.34	0.28	0.22	0.30
ILMN_2187718	COX17	0.29	0.28	0.31	0.36
ILMN_2128128	SHFM1	0.22	0.27	0.41	0.43
ILMN_1799030	CMTM2	0.34	0.29	0.24	0.30
ILMN_1703538	AIF1	0.23	0.28	0.41	0.43
ILMN_2166865	ENY2	0.46	0.27	0.095	0.2
ILMN_1732328	LOC646200	0.40	0.25	0.12	0.21
ILMN_1680314	TXN	0.53	0.28	0.058	0.15
ILMN_1726239	TBCA	0.51	0.26	0.054	0.15
ILMN_2232936	UQCRH	0.53	0.27	0.048	0.15
ILMN_1746516	RPS25	0.42	0.26	0.11	0.21
ILMN_2189936	RPL36AL	0.071	0.26	0.79	0.79
ILMN_1791332	ATP5O	0.62	0.26	0.022	0.15
ILMN_1792528	LOC401206	0.50	0.26	0.055	0.15
ILMN_2189933	RPL36AL	0.31	0.26	0.24	0.30
ILMN_2225887	ATP5EP2	0.58	0.27	0.036	0.15
ILMN_1661945	C14orf156	0.62	0.27	0.027	0.15
ILMN_1745343	ZMAT2	0.54	0.27	0.045	0.15
ILMN_1680967	CIP29	0.51	0.28	0.07	0.16
ILMN_1652073	LOC653658	0.57	0.25	0.025	0.15

Six probes (UQCRH, ATP5O, ATP5EP2, C14orf156, ZMAT2, and LOC653658) were found to be nominally significantly ($p < 0.05$) associated with the use of AChEIs (Table 2). The box plots for these six probes are shown in Fig. 2. Only two of these—LOC653658 and C14orf156—were still nominally associated with medication use when presence or absence of *APOE ε4* (and its interaction with medication use) was also accounted for in the model. Visually there does not appear to be a significant difference in expression of these genes between groups.

Proteomic study

For the 168 AD subjects with proteomic, demographic, and medication data, no significant differences in gender, age, *APOE* status, or MMSE scores were observed between the AChEI group and the non-AChEI group ($p < 0.05$; Table 3). None of these variables were thus factored in as covariates for the proteomic analysis.

The four proteins analyzed showed no significant difference in quantity between both groups in a linear model, with none passing significance thresholds either at the nominal ($p < 0.05$) or multiple testing corrected ($q < 0.05$) threshold (see Table 4).

Table 3
Demographic information for subjects included in the proteomic analysis

	AChEI	Non-AChEI	*p*-value
Total Number	129	39	–
Males/Females	48/81	9/30	0.12
Median Age (IQR)	77 (10)	78 (9.5)	0.45
Median MMSE (IQR)	21 (8)	21 (8)	0.63
APOE ε4 Status (0/1/2)	55/48/26	19/18/2	0.074

AChEI, Group taking acetylcholinesterase inhibitors; Non-AChEI, Group not on any AD cognitive enhancing medication; MMSE, Mini-Mental State Examination.

DISCUSSION

The analysis conducted shows that none of the biomarkers studied are differentially expressed in subjects taking AChEIs. Our sample had reasonable statistical power to detect large effects on gene expression markers, and both medium or large effects on protein markers. This implies that AChEIs are not a large confounding factor affecting the most promising gene or protein biomarkers identified in the studies by Lunnon et al. [9] and Sattlecker et al. [12]. If AChEIs had been found to be a large confounding factor, it would have undermined the diagnostic/enrichment potential of the biomarkers identified in those studies. The results of this study, though small in scope and by

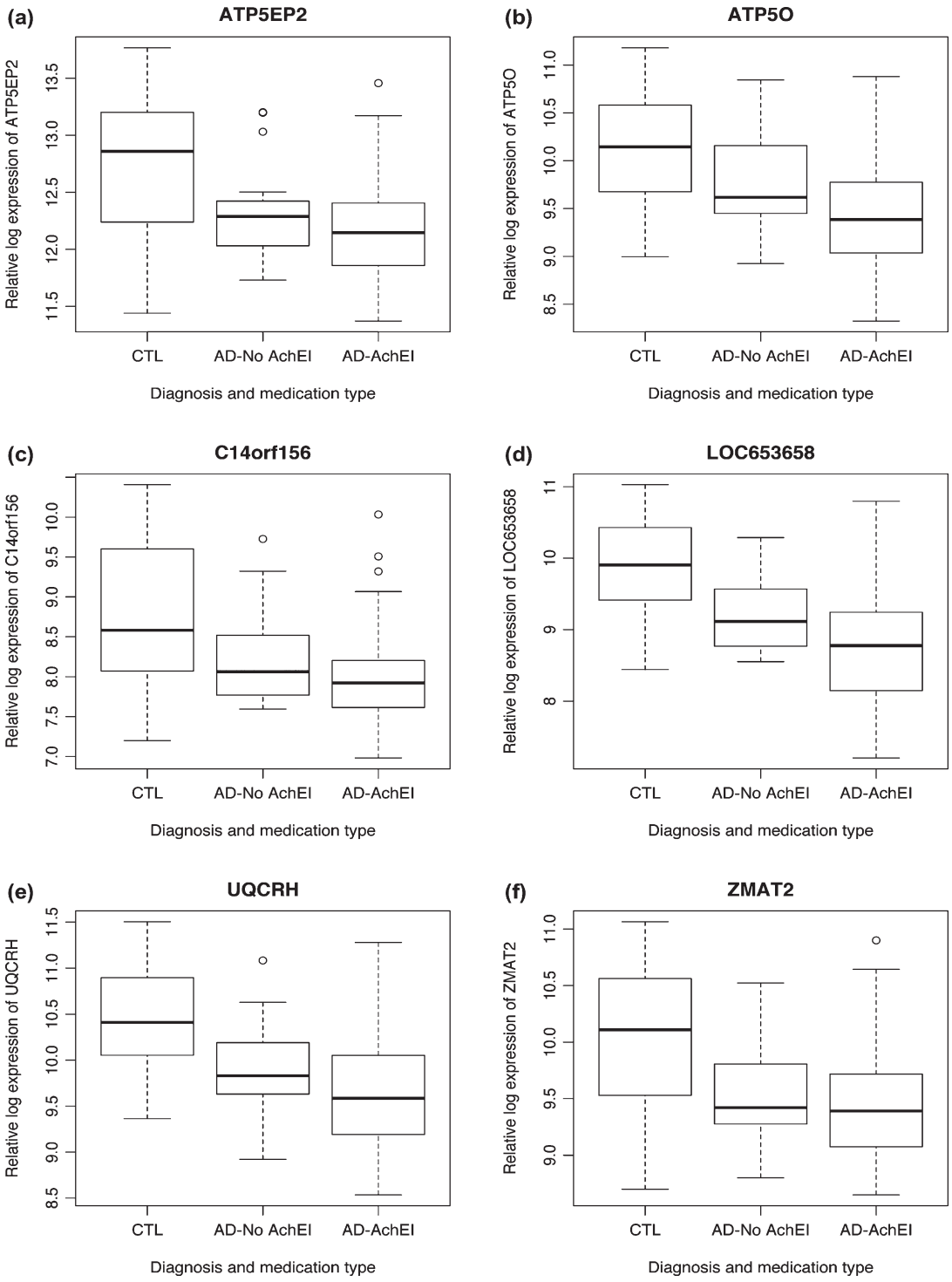


Fig. 2. Box plots of the six nominally significant genes found in the gene expression analysis. AD subjects on acetylcholinesterase inhibitors (AChEI) are compared with AD subjects not on AChEIs (Non-AChEI) and control subjects without AD (CTL). There do not appear to be a significant difference in gene expression between AD subjects with and without AChEIs.

Table 4
Results for the proteomic analysis showing the four proteins analyzed with details of the UniProt ID, coefficient value, standard error, *p*-value and q-value. *q-values are approximately equal, exact equality of these q-values is an artifact of the approximation method used by the R function 'p.adjust'

UniProt ID	Protein Name	Coefficient	Standard Error	*p*-value	q-value
P07288, P01011	PSA-ACT	–0.049	0.19	0.80	0.80*
Q9UGM5	Fetuin B	0.083	0.19	0.66	0.80*
P10909	Clusterin	0.12	0.19	0.54	0.80*
P01298	Pancreatic Prohormone	0.17	0.19	0.37	0.80*

PSA-ACT, prostate specific antigen complexed to α1-antichymotrypsin.

no means comprehensive, are therefore encouraging and strengthen the validity of these studies. As the first study to examine the possibility of AD blood biomarkers being confounded by AChEIs, it also highlights a previously neglected potential confounding variable.

In this study the sample size of both groups tested was notably imbalanced. The number of subjects in the group not on any AD medication was significantly smaller than the group on AChEIs in both the gene expression and proteomic analysis. This is understandable since most AD patients would be expected to have some sort of cognitive enhancing medication as treatment. It would also have been interesting to include an analysis of subjects using memantine, the other main class of AD cognitive drugs. Yet in this study too few subjects were on memantine to allow a thorough investigation of this. This limitation could be overcome should larger studies be conducted in the future.

While in this study AChEIs were considered as a single class of drug, three separate AChEIs (donepezil, galantamine, and rivastigmine) are the mainstay in AD therapy [20]. Though all three drugs work similarly by inhibiting acetylcholinesterase and preventing the breakdown of acetylcholine at the post-synaptic cleft, it is likely that that there are subtle differences in their effects on underlying biological processes. Therefore if an adequate sample size is available in future studies, all three drugs should be studied individually for their effects on AD blood biomarkers. This is pertinent since studies revealing differences in gene expression as a result of AChEI use have studied AChEIs individually. Specifically, Reale et al. [23] investigated the effects of donepezil on blood inflammatory markers in AD patients, while Andin et al. [38] investigated the effect of rivastigmine on the glutamate transporter rEAAC1 blood mRNA expression in mice models, both revealing significant effects by the drugs on gene expression.

One possible way of improving the sample size in future studies is to ensure the collection of appropriate medication information in any research cohort, as some subjects were excluded from our study due to incomplete demographic data. Furthermore with hindsight, medication could have been included as a covariate in the discovery stage of biomarker studies, instead of analyzed *post-hoc*. We have not seen this approach applied in any of the discovery studies to date and this should therefore be considered for future biomarker studies.

Besides drugs that improve cognitive symptoms, many AD patients are also on other medications to manage the non-cognitive symptoms of the disease. People with dementia are far more susceptible to psychiatric conditions such as mood disorders and psychosis [39]. In such cases, drugs such as antidepressants, neuroleptics, sedatives, and hypnotics are often required for treatment [19, 40]. Citalopram, an antidepressant, has been found to affect gene expression in AD lymphocytes [41]. Thus it is possible that other psychotropic medications could prove to be a confounding factor in blood biomarker studies and should be explored in future studies. Given that each subject could potentially be on multiple medications, it may also be possible to conduct multivariate analyses to determine if a combination of medications would yield a significant difference in gene expression, although larger cohorts would be needed to identify this.

This study considered promising biomarkers individually, looking at whether the expression of single genes or proteins are affected by AChEIs. However, Sattlecker et al. [12] and Lunnon et al. [42], as well as other researchers, have also proposed multivariate biomarker models of AD. These multivariate models could also be investigated to determine if AChEIs, or any other medications, affect their potential utility.

It is also important to consider the possible effects of medication on other potential blood biomarkers. For example, many other potential biomarkers of AD have been highlighted in recent reviews of blood gene expression [17] and protein levels [16].

CONCLUSION

Overall this exploratory study has found no evidence that AChEIs are a large confounding factor

for the most promising AD blood-based biomarkers identified in both studies. This gives an encouraging indication that the use of AChEIs is unlikely to affect the validity of these biomarkers in potential diagnosis or enrichment applications. However, more comprehensive studies need to be conducted to explore the full effects of AChEIs on these and other proposed blood-based biomarkers of AD.

ACKNOWLEDGMENTS

This work was supported by the Alzheimer's Society (NV), InnoMed (Innovative Medicines in Europe), an integrated project funded by the European Union of the Sixth Framework program priority (FP6-2004-LIFESCIHEALTH-5); the Alzheimers Research Trust UK; the John and Lucille van Geest Foundation; and the NIHR Biomedical Research Centre for Mental Health and Biomedical Research Unit for Dementia at the South London, Maudsley NHS Foundation Trust and Kings College London, and a joint infrastructure grant from Guy's and St Thomas' Charity and the Maudsley Charity; Kuopio University Hospital (HS) and funding from UEF- BRAIN (HS). Steven Kiddle is supported by an MRC Career Development Award in Biostatistics (MR/L011859/1). The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement number 115372, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution. SOMAscan and SOMAmer are trademarks of SomaLogic, Inc. The authors acknowledge Abhishek Dixit for his help with data management and access. We would also like to thank peer reviewers for helpful advice.

Authors' disclosures available online (http://j-alz.com/manuscript-disclosures/15-0289r1).

REFERENCES

[1] Burns A, Iliffe S (2009) Alzheimer's disease. *BMJ* **338**, b158.

[2] Frankfort SV, Tulner LR, van Campen JP, Verbeek MM, Jansen RW, Beijnen JH (2008) Amyloid beta protein and tau in cerebrospinal fluid and plasma as biomarkers for dementia: A review of recent literature. *Curr Clin Pharmacol* **3**, 123-131.

[3] Frisoni GB, Fox NC, Jack CR Jr, Scheltens P, Thompson PM (2010) The clinical use of structural MRI in Alzheimer disease. *Nat Rev Neurol* **6**, 67-77.

[4] Hampel H, Lista S, Khachaturian ZS (2012) Development of biomarkers to chart all Alzheimer's disease stages: The royal road to cutting the therapeutic Gordian Knot. *Alzheimers Dement* **8**, 312-336.

[5] Ho L, Fivecoat H, Wang J, Pasinetti GM (2010) Alzheimer's disease biomarker discovery in symptomatic and asymptomatic patients: Experimental approaches and future clinical applications. *Exp Gerontol* **45**, 15-22.

[6] Villemagne VL, Burnham S, Bourgeat P, Brown B, Ellis KA, Salvado O, Szoeke C, Macaulay SL, Martins R, Maruff P, Ames D, Rowe CC, Masters CL, Australian Imaging Biomarkers, Lifestyle (AIBL) Research Group (2013) Amyloid beta deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: A prospective cohort study. *Lancet Neurol* **12**, 357-367.

[7] Booij BB, Lindahl T, Wetterberg P, Skaane NV, Saebo S, Feten G, Rye PD, Kristiansen LI, Hagen N, Jensen M, Bardsen K, Winblad B, Sharma P, Lonneborg A (2011) A gene expression pattern in blood for the early detection of Alzheimer's disease. *J Alzheimers Dis* **23**, 109-119.

[8] Fehlbaum-Beurdeley P, Jarrige-Le Prado AC, Pallares D, Carriere J, Guihal C, Soucaille C, Rouet F, Drouin D, Sol O, Jordan H, Wu D, Lei L, Einstein R, Schweighoffer F, Bracco L (2010) Toward an Alzheimer's disease diagnosis via high-resolution blood gene expression. *Alzheimers Dement* **6**, 25-38.

[9] Lunnon K, Ibrahim Z, Proitsi P, Lourdusamy A, Newhouse S, Sattlecker M, Furney S, Saleem M, Soininen H, Kloszewska I, Mecocci P, Tsolaki M, Vellas B, Coppola G, Geschwind D, Simmons A, Lovestone S, Dobson R, Hodges A, AddNeuroMed Consortium (2012) Mitochondrial dysfunction and immune activation are detectable in early Alzheimer's disease blood. *J Alzheimers Dis* **30**, 685-710.

[10] Ray S, Britschgi M, Herbert C, Takeda-Uchimura Y, Boxer A, Blennow K, Friedman LF, Galasko DR, Jutel M, Karydas A, Kaye JA, Leszek J, Miller BL, Minthon L, Quinn JF, Rabinovici GD, Robinson WH, Sabbagh MN, So YT, Sparks DL, Tabaton M, Tinklenberg J, Yesavage JA, Tibshirani R, Wyss-Coray T (2007) Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins. *Nat Med* **13**, 1359-1362.

[11] Doecke JD, Laws SM, Faux NG, Wilson W, Burnham SC, Lam CP, Mondal A, Bedo J, Bush AI, Brown B, De Ruyck K, Ellis KA, Fowler C, Gupta VB, Head R, Macaulay SL, Pertile K, Rowe CC, Rembach A, Rodrigues M, Rumble R, Szoeke C, Taddei K, Taddei T, Trounson B, Ames D, Masters CL, Martins RN, Alzheimer's Disease Neuroimaging Initiative, Australian Imaging Biomarker Lifestyle Research, Group (2012) Blood-based protein biomarkers for diagnosis of Alzheimer disease. *Arch Neurol* **69**, 1318-1325.

[12] Sattlecker M, Kiddle SJ, Newhouse S, Proitsi P, Nelson S, Williams S, Johnston C, Killick R, Simmons A, Westman E, Hodges A, Soininen H, Kloszewska I, Mecocci P, Tsolaki M, Vellas B, Lovestone S, the AddNeuroMed Consortium, Dobson RJ (2014) Alzheimer's disease biomarker discovery using SOMAscan multiplexed protein technology. *Alzheimers Dement* **10**, 724-734.

[13] Zurbig P, Jahn H (2012) Use of proteomic methods in the analysis of human body fluids in Alzheimer research. *Electrophoresis* **33**, 3617-3630.

[14] Lista S, Faltraco F, Prvulovic D, Hampel H (2013) Blood and plasma-based proteomic biomarker research in Alzheimer's disease. *Prog Neurobiol* **101-102** 1-17.

[15] Kiddle SJ, Sattlecker M, Proitsi P, Simmons A, Westman E, Bazenet C, Nelson SK, Williams S, Hodges A, Johnston C, Soininen H, Kloszewska I, Mecocci P, Tsolaki M,

Vellas B, Newhouse S, Lovestone S, Dobson RJ (2014) Candidate blood proteome markers of Alzheimer's disease onset and progression: A systematic review and replication study. *J Alzheimers Dis* **38**, 515-531.

[16] Chiam JT, Dobson RJ, Kiddle SJ, Sattlecker M (2014) Are blood-based protein biomarkers for Alzheimer's disease also involved in other brain disorders? A systematic review. *J Alzheimers Dis* **43**, 303-314.

[17] Han G, Wang J, Zeng F, Feng X, Yu J, Cao HY, Yi X, Zhou H, Jin LW, Duan Y, Wang YJ, Lei H (2013) Characteristic transformation of blood transcriptome in Alzheimer's disease. *J Alzheimers Dis* **35**, 373-386.

[18] Henriksen K, O'Bryant SE, Hampel H, Trojanowski JQ, Montine TJ, Jeromin A, Blennow K, Lönneborg A, Wyss-Coray T, Soares H, Bazenet C, Sjögren M, Hu W, Lovestone S, Karsdal MA, Weiner MW (2014) The future of blood-based biomarkers for Alzheimer's disease. *Alzheimers Dement* **10**, 115-131.

[19] Lauterbach EC (2012) Psychotropic drug effects on gene transcriptomics relevant to Alzheimer disease. *Alzheimer Dis Assoc Disord* **26**, 1-7.

[20] O'Brien JT, Burns A, BAP Dementia Consensus Group (2011) Clinical practice with anti-dementia drugs: A revised (second) consensus statement from the British Association for Psychopharmacology. *J Psychopharmacol* **25**, 997-1019.

[21] Lane RM, Kivipelto M, Greig NH (2004) Acetylcholinesterase and its inhibition in Alzheimer disease. *Clin Neuropharmacol* **27**, 141-149.

[22] Singh M, Kaur M, Kukreja H, Chugh R, Silakari O, Singh D (2013) Acetylcholinesterase inhibitors as Alzheimer therapy: From nerve toxins to neuroprotection. *Eur J Med Chem* **70**, 165-188.

[23] Reale M, Iarlori C, Gambi F, Feliciani C, Salone A, Toma L, DeLuca G, Salvatore M, Conti P, Gambi D (2004) Treatment with an acetylcholinesterase inhibitor in Alzheimer patients modulates the expression and production of the pro-inflammatory and anti-inflammatory cytokines. *J Neuroimmunol* **148**, 162-171.

[24] Lovestone S, Francis P, Strandgaard K (2007) Biomarkers for disease modification trials–the innovative medicines initiative and AddNeuroMed. *J Nutr Health Aging* **11**, 359-361.

[25] Lovestone S, Francis P, Kloszewska I, Mecocci P, Simmons A, Soininen H, Spenger C, Tsolaki M, Vellas B, Wahlund LO, Ward M, AddNeuroMed Consortium (2009) AddNeuroMed–the European collaboration for the discovery of novel biomarkers for Alzheimer's disease. *Ann N Y Acad Sci* **1180**, 36-46.

[26] Lovestone S, Thambisetty M (2009) Biomarkers for Alzheimer's disease trials–biomarkers for what? A discussion paper. *J Nutr Health Aging* **13**, 334-336.

[27] McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM (1984) Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* **34**, 939-944.

[28] Folstein MF, Folstein SE, McHugh PR (1975) Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* **12**, 189-198.

[29] Reisberg B, Ferris SH, de Leon MJ, Crook T (1988) Global Deterioration Scale (GDS). *Psychopharmacol Bull* **24**, 661-663.

[30] Rosen WG, Mohs RC, Davis KL (1984) A new rating scale for Alzheimer's disease. *Am J Psychiatry* **141**, 1356-1364.

[31] Dixit A, Dobson R (2014) CohortExplorer: A generic application programming interface (API) for entity attribute value database schemas. *JMIR Med Inform* **2**, e32.

[32] Du P, Kibbe WA, Lin SM (2008) lumi: A pipeline for processing Illumina microarray. *Bioinformatics* **24**, 1547-1548.

[33] Ding LH, Xie Y, Park S, Xiao G, Story MD (2008) Enhanced identification and biological validation of differential gene expression via Illumina whole-genome expression arrays through the use of the model-based background correction methodology. *Nucleic Acids Res* **36**, e58.

[34] Oldham MC, Langfelder P, Horvath S (2012) Network methods for describing sample relationships in genomic datasets: Application to Huntington's disease. *BMC Syst Biol* **6**, 63-0509-6-63.

[35] Johnson WE, Li C, Rabinovic A (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118-127.

[36] Gold L, Ayers D, Bertino J, Bock C, Bock A, Brody EN, Carter J, Dalby AB, Eaton BE, Fitzwater T, Flather D, Forbes A, Foreman T, Fowler C, Gawande B, Goss M, Gunn M, Gupta S, Halladay D, Heil J, Heilig J, Hicke B, Husar G, Janjic N, Jarvis T, Jennings S, Katilius E, Keeney TR, Kim N, Koch TH, Kraemer S, Kroiss L, Le N, Levine D, Lindsey W, Lollo B, Mayfield W, Mehan M, Mehler R, Nelson SK, Nelson M, Nieuwlandt D, Nikrad M, Ochsner U, Ostroff RM, Otis M, Parker T, Pietrasiewicz S, Resnicow DI, Rohloff J, Sanders G, Sattin S, Schneider D, Singer B, Stanton M, Sterkel A, Stewart A, Stratford S, Vaught JD, Vrkljan M, Walker JJ, Watrobka M, Waugh S, Weiss A, Wilcox SK, Wolfson A, Wolk SK, Zhang C, Zichi D (2010) Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS One* **5**, e15004.

[37] Cohen J (1988) *Statistical Power Analysis for the Behavioral Sciences*, Lawrence Erlbaum Associates, Hillsdale, New Jersey.

[38] Andin J, Enz A, Gentsch C, Marcusson J (2005) Rivastigmine as a modulator of the neuronal glutamate transporter rEAAC1 mRNA expression. *Dement Geriatr Cogn Disord* **19**, 18-23.

[39] Robert PH, Verhey FR, Byrne EJ, Hurt C, De Deyn PP, Nobili F, Riello R, Rodriguez G, Frisoni GB, Tsolaki M, Kyriazopoulou N, Bullock R, Burns A, Vellas B (2005) Grouping for behavioral and psychological symptoms in dementia: Clinical and biological aspects. Consensus paper of the European Alzheimer disease consortium. *Eur Psychiatry* **20**, 490-496.

[40] National Collaborating Centre for Mental Health (UK) (2007) Dementia: A NICE-SCIE Guideline on Supporting People With Dementia and Their Carers in Health and Social Care. *National Institute for Health and Clinical Excellence: Guidance*, British Psychological Society, Leicester, UK.

[41] Palotas A, Puskas LG, Kitajka K, Palotas M, Molnar J, Pakaski M, Janka Z, Penke B, Kalman J (2004) The effect of citalopram on gene expression profile of Alzheimer lymphocytes. *Neurochem Res* **29**, 1563-1570.

[42] Lunnon K, Sattlecker M, Furney SJ, Coppola G, Simmons A, Proitsi P, Lupton MK, Lourdusamy A, Johnston C, Soininen H, Kloszewska I, Mecocci P, Tsolaki M, Vellas B, Geschwind D, Lovestone S, Dobson R, Hodges A, dNeuroMed Consortium (2013) A blood gene expression marker of early Alzheimer's disease. *J Alzheimers Dis* **33**, 737-753.

Appendix B

Statistical Methods in epidemiology: Bio-informaticians perspective *Practical Psychiatric Epidemiology, Second Edition (in press)*

Statistical Methods in epidemiology: Bio-informatician's perspective

Nicola Voyle, Maximilian Kerz, Steven Kiddle and Richard Dobson

1. Introduction

This chapter walks through the basic statistical methods needed to explore, analyse and interpret data collected in epidemiological studies. We begin with raw data, highlighting essential data visualisation and cleaning steps before providing intuitive explanations for some of the most commonly used modelling tools from statistics and bioinformatics. We end by outlining the importance of testing results in independent data sets, where possible, and suggest other methodologies where this is not possible. After studying this chapter you should have a comprehensive idea of how to perform data analysis and have identified areas that you wish to research further to gain a more technical understanding.

2. Data collection

The first stage of any analysis is to collect data from a well-designed study. We begin by looking at the raw data that arises from epidemiological studies and highlighting the first stages in creating clean data that can be used to draw informative conclusions through analysis. We move on to describe big data and associated challenges before emphasising the importance of reproducibility in research.

2.1 Data formats

Raw data can come in many different types and formats, ranging from hand-written lab sheets and unformatted excel files, to images. In order to analyse and explore data it is vital to fully understand the structure and content of your raw data set. Depending on its nature, and the type of analyses you would like to run, an initial pre-processing step is recommended or even unavoidable. Examples of pre-processing steps include transforming qualitative data into machine-readable categories or using an image analysis pipeline to extract important features of an image and cast them into a structured, tabular format.

2.2 Data exploration

The aim of data exploration is to use descriptive methods to get a statistical and visual impression of the data (Sayad, 2011). Exploratory statistics allow quantification and visualisation of qualitative variables, such as categorical data, as well as continuous variables. Perhaps most importantly, data exploration is the first step in identifying new patterns and/or associations. Data exploration is of particular importance in large data samples where a preliminary impression of the data helps to identify potential

confounding factors, anomalies and missing data (Podsakoff et al., 2003).

2.3 Data cleaning

Data collection and aggregation into a master data set can result in irregularities and missingness. Human error during data collection or participant-dropouts during a longitudinal study are examples that can result in the former. The process of removing or accounting for these impurities within the data set is known as data cleaning.

The extent of data cleaning depends on the specific method that you are aiming to apply and ranges from basic methods to advanced pre-processing. Amongst the most common mistakes that lead to impure data sets is failure to look at the data at all, or in tabular form only. Basic data exploration and plots can easily identify missing values, label switching and unit mismatches. A further advantage of exploratory plots is an indication of skewness that can reveal biases within the data (Kuhn and Johnson, 2013). These can then be accounted for before applying additional analyses.

2.4 Missing data

Missingness in statistics is defined as the lack of data for a variable in an observation. It is a common phenomenon in the realm of data analysis and is grouped into three categories: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR), also known as informative missingness (van Buuren, 2012).

- MCAR: the probability of being missing is equal for all observations. In other words, the cause for missingness is unrelated to the data.

- MAR: the probability of being missing is equal for a particular observation only.
- MNAR: the probability of being missing varies depending on another observation.

An example of MNAR would be an increasing dropout rate in a clinical trial based on the administered concentration of an experimental drug. Even though the data exhibits increasingly more missingness, it is highly informative on drug adherence in patients (Kuhn and Johnson, 2013).

Treating missingness is a challenging topic that can range from simple data removal to application of prediction models for imputation. Depending on the nature of the data, as well as the type and extent of missingness, different methods should be applied (Schafer and Graham, 2002). The exact intricacies of treating missingness in data are beyond the scope of this chapter.

2.5 Big data

Big data is vaguely defined. However, the most commonly cited definition of big data is by Laney (2001) from Gartner as “high volume, velocity and variety information that demand cost-effective, innovative forms of information processing for enhanced insight and decision making”. Additionally, variability within a data type is an important characteristic of big data and should be added to the definition. There is no clear boundary to when conventional data turn into big data, however any combination of complexity, size, and advanced analytics technology satisfies the definition by Laney (2001). In fact, data sets do not even need to be particularly big, just complex, to be considered big data.

2.6 Reproducibility

Reproducibility allows re-running an analysis and always obtaining the same result when applied to the same data set. This is particularly important for the validity and generalisability of your analysis should you want to publish your results. Unfortunately, reproducing complex analyses is associated with challenges, such as software updates that change the outcome of an analysis, incomplete or non-existent documentation, and failing to record intermediary results.

Detailed documentation and scripted analyses are a great way to ensure that your research is reproducible. This will not only benefit you but also the community of psychiatric epidemiology.

2.7 Classification vs. regression

It is likely that the data collection and cleaning covered above has been to answer a specific research question. Before moving forward it is important to identify whether the question you are answering is a classification or regression question:

- Classification: Allocation of an observation to one of several groups. For example, classifying a brain scan as positive or negative for a defined pathology.
- Regression: Predicting a continuous outcome for an observation. For example, predicting hippocampal volume from an MRI scan.

Throughout this chapter we will refer to classification however, most methods can also be applied in a regression setting.

3. Feature identification and selection

Once a data set has been processed (if not before) an analysis plan needs to be created. Conventional statistical methods (not covered in this chapter) require the number of variables in the data set to be less than the number of samples. However, machine learning algorithms do not have such a limitation. Machine learning algorithms use induction to learn; they aim to improve prediction based on the data provided. Although they can handle many thousands of predictors it is often computationally expensive to do so. Additionally, models with many predictors are very hard to interpret. Therefore, we aim to reduce the number of variables we use in modeling via feature selection. Feature selection removes redundancy and noise from data leaving only the most discriminatory features while reducing problems of dimensionality. Guyon and Elisseeff (2003) provide a good introduction to feature selection techniques.

4 Method selection

After selecting features to include in the analysis, or planning the feature selection techniques to use alongside modeling, the modeling algorithm must be selected. The choice of algorithm is important and should be tailored to the question you are looking to answer and the data you have. Do you have more predictor variables than samples? Some methods (for example support vector machines (SVMs)) are good at coping with high dimensionality datasets with a small number of samples, however this may require large amounts of computer memory. Do you have balanced training data? Some methods, such as SVMs and decision trees, are sensitive to imbalance in the training dataset and should therefore be avoided if you have unbalanced data. Where possible it is preferable to test a number of classifiers to identify the most appropriate choice for the specific problem.

4.1 Supervised vs. unsupervised machine learning

The most common types of machine learning algorithms are supervised learning and unsupervised learning. Supervised learning takes a set of examples that are labelled and creates a set of rules to classify samples where the status is unknown. In contrast, unsupervised learning models a set of inputs where labelled examples are not available. The algorithm finds a way of clustering the data based upon the known features and then provides descriptions for these clusters.

In this section we will introduce some vital concepts before discussing some supervised and unsupervised algorithms that are amongst the most commonly used within the field of bioinformatics.

Box 1: Supervised learning example

Predicting whether a non-synonymous SNP (nsSNP) is disease related is a question that can be addressed via machine learning methods. Supervised learning is appropriate because the aim is to assign an nsSNP to one of a number of classes. It is possible to use a set of nsSNPs where the disease status is known as a training set to form a set of rules. These rules can then be used to make a prediction for nsSNPs where the function is unknown.

4.2 Important concepts

4.2.1 Training, testing and validation

A common problem of supervised learning methods is over-fitting. An algorithm can be trained for too long on one particular data set so that it fails to generalise the information learned to similar data sets. A number of methods are available for evaluating machine learning results and showing the results are general enough to be applied to other data (Hand *et al.*, 2001). To successfully train a supervised learning algorithm, one should aim to generate three data sets from the original data set.

- A training set: used to train the algorithm.
- A validation set: to track how well the algorithm is generalizing and to perform parameter tuning.
- A test set: an unseen data set on which the finalised algorithm's performance is tested.

As machine learning algorithms require substantial training data, the usual

distribution of the training, validation and testing sets is 2:1:1. However, in the case of limited data availability, different cross-validation (CV) methods can be employed.

4.2.2 Cross validation

In cross-validation the data is divided into a number (n) of 'folds'. Each fold is treated as the validation dataset in turn, with the remaining n-1 folds being used as training data. CV is especially useful for smaller datasets (Kohavi, 1995). The performance of the classifier on each fold is measured and then a final accuracy is calculated based upon the average of all n folds.

4.2.3 Bootstrap

Sampling from a whole dataset, with replacement, creates a bootstrap sample. The bootstrap sample is the same size as the original dataset and contains, on average, one third of all observations in the whole dataset. The observations not included in the bootstrap sample are called the out of bag (OOB) data. As sampling with replacement creates the bootstrap sample, it can contain repeated instances (Efron & Tibshirani, 1994).

Bootstrapping is a useful tool to maintain the size of your training dataset while also creating validation data. The performance of the classifier is calculated by averaging the accuracy from the OOB data of each bootstrap sample.

4.2.4 Hyper-plane

A hyper-plane is a subspace with one less dimension than the whole space under consideration. This is easiest to imagine in small dimensions. For example, if we

consider a problem in two dimensions (for example, plotting age against height) a hyper-plane is any one-dimensional line running through the space. Similarly, if we consider a problem in three dimensions (for example, plotting age against height and weight) a hyper-plane is any two-dimensional plane.

4.3 Unsupervised methods

Unsupervised learning has a long and distinguished history within modern epidemiology. For example, John Snow a physician in the mid-19th century is famous for discovering the source of Cholera outbreaks, which were at the time believed to be due to 'bad air'. To achieve this he used an unsupervised data analysis method. During the 1854 Cholera outburst in Soho, London, John Snow recorded the location of deaths onto a map of Soho, showing that deaths were clustered around an intersection of Broad Street. He later identified the source as an infected water pump (Johnson, 2006).

Unsupervised machine learning extends the concept of looking for patterns to complex multidimensional datasets. Like John Snow's map it seeks to identify patterns within datasets. Patterns extracted from data using unsupervised machine learning can also be used later in supervised approaches to identify the underlying causes of various patterns, analogous to John Snow adding information on water pumps onto his maps.

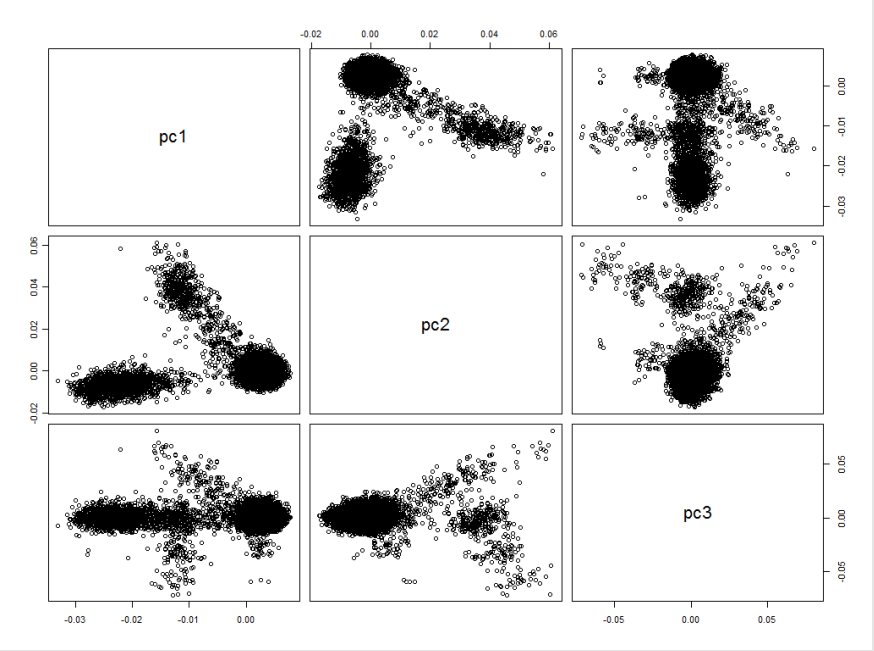
4.3.1 Principal components analysis (PCA)

A commonly used unsupervised approach to data analysis is Principal Components Analysis (PCA). For example, PCA is often used in genetic epidemiology to take into account population stratification and systematic differences in allele frequencies that are often due to ancestry (Price *et al.*, 2006). It can also be invaluable for visualizing outliers (Price *et al.,* 2006) and batch effects (Alter *et al*., 2000) in high-dimensional datasets, for example whole genome gene expression data. As such, it is often used as part of the important 'quality control' process.

PCA takes a set of possibly correlated variables and returns a set of Principal Components (PCs), which are orthogonal hyper-planes that explain maximal variance in the predictors. A useful feature of the PCs is that they are given in rank order for the proportion of total variation they explain; PC 1 explains more of the variation in a dataset than PC 2 does. A useful consequence of this is that it is often possible to summarise a large proportion of the total variance using only 2 or 3 principal components allowing high-dimensional datasets to be summarised in 2D or 3D plots. For example Figure 1 shows the effect of ancestry on PCA applied to genetic data. The proportion of variance that each PC explains can be visualised in a 'scree plot', helping you to choose the number of PCs to use for plotting or further analysis. More details on PCA are reviewed in Jolliffe (2014).

Box 2: Principal components analysis example

Figure 1 – Principal Components Analysis of 125 Ancestry Informative Markers (AIMs). Scatterplots are shown comparing the first three Principal Components. Individuals are represented by one circle in each scatterplot. The first 3 Principal Components summarise the majority of the variability in ancestry, and help to distinguish individuals of European, African-American, Hispanic and Asian descent. Figure generated by Dr William S Bush, Assistant Professor at Case Western Reserve University, Cleveland Ohio.



We are grateful to Dr Bush for permission to reproduce this figure, which originally appeared at his blog http://www.gettinggeneticsdone.com/2011/10/new-dimension-to-principal-components_27.html

4.3.2 Clustering

Methods such as PCA can be used to summarise variability in a dataset, often showing that data points can be divided into 'groups' or 'clusters'. The general definition of a cluster is that objects within a cluster are more similar to each other than they are to objects outside of that cluster. But how do you automate the assignment of data points to clusters? And how many clusters exist?

The problem of assigning data points to clusters is an unsupervised machine learning problem, for which many different approaches exist. One of the most commonly used approaches is hierarchical clustering, which generates a hierarchy tree with a data point lying at the end of every branch (Ward, 1963). The branches for the most similar data points are connected first, and this is repeated until eventually all data points are connected. Data points are separated into clusters by applying a threshold to the tree. The problem with this approach is that as branches are combined one at a time, the final grouping may not always be optimal.

A common alternative to hierarchical clustering is called K-means (Jain, 2010). This involves looking for K clusters in the data, where K is a number. The K-means procedure proceeds as follows:

- Data points are randomly assigned into K different clusters.
- The mean of each cluster is calculated.
- Each data point is assigned to belong to the cluster whose mean it is closest to.
- Steps 2 – 3 are repeated until an acceptable solution is reached.

K-means solutions are often more optimal than a hierarchical clustering solution. However, to perform K-means, a 'mean' must be definable for your dataset. Where this is not possible, alternatives such as K-centres or affinity propagation can be used.

For both approaches, you need to specify a threshold or a number of clusters in advance. A commonly used approach to find the optimal number or threshold is the silhouette plot, which provides a visualization of the quality of clusters (Rousseeuw, 1987).

4.4 Supervised methods

4.4.1 K nearest neighbours (KNN)

K nearest neighbours (KNN) is arguably the simplest supervised learning method (Altman, 1992). The basic idea is that a new observation is classified to the modal class of the k closest observations from the training data (where k is some integer). There are a number of methods to define what we mean by 'closest' but these are not discussed here. The example given in Box 3 should clarify this idea. In order to perform an analysis using this method we must choose a suitable value for k. This is known as parameter tuning and is usually performed by using several rounds of cross-validation to choose the value that gives the best results. If k is too low the model can over-fit to the training data and conversely, if k is too high, the model may under-fit. Although KNN provides a method with very simple intuition the model fit can be hard to interpret; it is solely based on training data so there is no description of the model. Furthermore, it can become computationally expensive on large datasets.

Box 3: KNN example

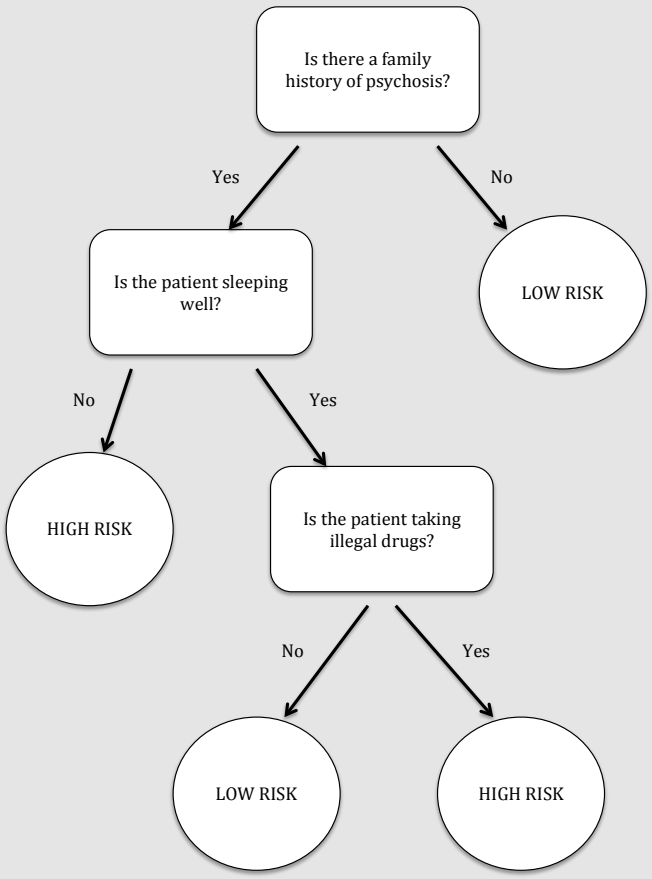
We have a group of patients presenting with memory impairment; they have been classified as having either dementia or short-term memory loss. Suppose that for each of them we have only two pieces of clinical information: age and hippocampal volume from an MRI scan. We could plot this information and colour code by the subject's classification. When a new patient enters the clinic presenting with memory impairment and with these two measures available we can add them to the plot. We can consider the 5 patients who are closest to the new patient on the graph and output the most common class from these 5 subjects as the classification.

4.4.2 Decision tress

Decision trees are supervised classifiers composed of a graph (tree structure) of decisions (Quinlan, 1993). Each point in the tree where a decision is made is called a node and the result of each decision determines which branch to follow. The decisions are usually simple single attribute tests to divide the data. A leaf represents the predicted class based on values at the nodes on the path from the root (the first decision point). Decision trees have an advantage over many classifiers in that they produce interpretable rules. Once a tree has been built new instances can be classified by starting at the root and following a path down to a leaf. An example of a decision tree can be seen in Figure 2 where a patient is classified as being at high or low risk of psychosis based on a number of attributes.

Box 4: Decision tree example

Figure 2: An example of a decision tree used to classify a patient as being at high or low risk of psychosis. The diagram shows decisions at the nodes and final classification at the leaves.



When the attribute at a node is categorical, there will be one branch for each attribute value. If the attribute is continuous, a decision will be made based on whether the instance is above or below a specific cut-off value. There are a number of methods for deciding which attribute should be used at each node but they are not discussed here.

A decision tree is complete when some criterion about the terminal nodes has been met, for example, when all the terminal nodes contain samples from only one class.

At this point the tree is often 'pruned' to prevent it being too specific to the training dataset. The aim is to produce a tree that is general enough to be applied to any new instances that require classification, avoiding over-fitting. The algorithms are efficient and therefore able to handle large volumes of data. However, one drawback to this approach is that the partitioning can cause interesting relationships between attributes to be lost.

4.4.3 Random forests (RF)

Random forest (RF) is a supervised classifier consisting of multiple decision trees (Breiman, 2001). The final class assigned to an observation is the modal class selected by the multiple decision trees. RF combines two machine learning methods: bootstrap sampling and random feature selection. Each tree is created from a bootstrap sample of the training data. OOB data is used to obtain an unbiased estimate of the error during the training. However, rather than using all features, RF randomly selects a subset of input variables to decide what decision should be made at each node of the tree. Advantages of random forest classifiers include the fact that the error can be balanced when the class population sizes are imbalanced and over-fitting can be avoided. There are also good methods available for handling missing data.

4.4.4 Partial least squares (PLS)

Partial least squares (PLS) modelling is very similar to PCA. Where PCA looks for orthogonal hyper-planes that explain maximal variance in the predictors, PLS looks to explain maximal covariance between the predictors and the outcome (Wold, 2004). It is this reliance on the outcome measure that means PLS modelling is supervised and is particularly suited for prediction problems. PLS modelling is useful when the

number of predictors is greater than the number of samples as it reduces the size of the predictor space by creating components.

4.4.5 Support vector machines (SVMs)

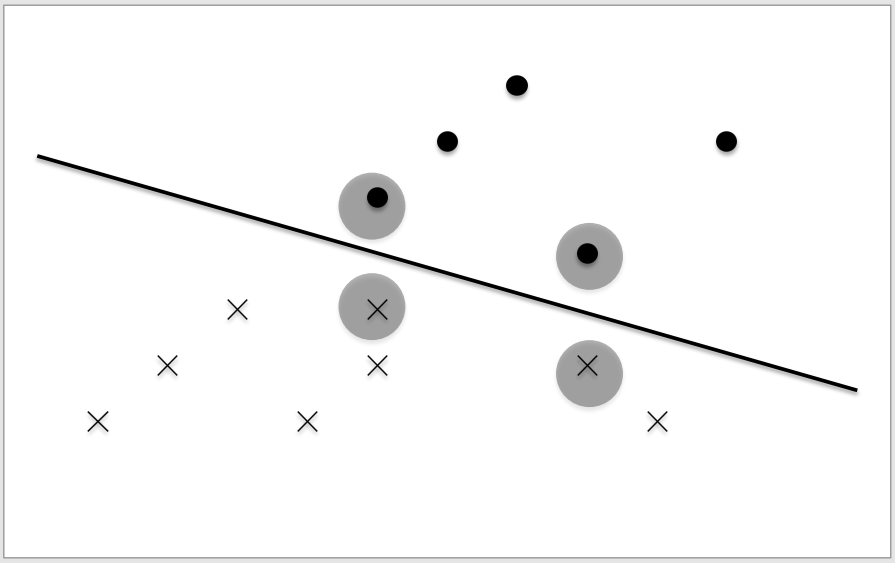
Support vector machines (SVMs) are a supervised learning classifier developed by Corte & Vapnik (1995). They have been shown to be very accurate in many disciplines including bioinformatics, benefitting from the ability to handle high dimensional data with a small number of instances, finding a good balance between training set accuracy and test data error. For a given set of training vectors labelled with two classes, a SVM can find the optimal linear hyper-plane that maximizes the margin between the two classes. An example of SVM classification in two dimensions is given in Figure 3.

SVMs can be extended to provide non-linear classification through the application of a kernel function and to multi-class classification but that is not discussed here.

Box 5: Support vector machine example

Figure 3: An example of a SVM classifying between two classes (circles and crosses).

The points acting as support vectors are highlighted.



5 Training a classifier

The process of building a model with initial data is called training.

An important point to consider when training a classifier is the presence (or lack thereof) of balanced data. The number of instances belonging to each class in the training set may be imbalanced resulting in a danger that the classifier will have a preference for selecting the most populated class because the classifier assumes that there is a greater chance of an instance belonging to this class (Barandela *et al.*, 2003). The result is that performance is reduced for the minority dataset. However, it may be the case, such as when detecting fraudulent telephone calls for example, that detecting the minority case is of greater importance (Fawcett & Provost, 1997).

6 Drawing conclusions from modelling

It is necessary to summarise the results of model testing using informative metrics. These metrics should quantify how well a model fits the test data.

6.1 Accuracy, sensitivity and specificity

There are four possible scenarios that could arise when classifying a new observation using a trained classifier. In this case we consider the example of a diagnostic test.

- A person with the disease is correctly classified as having the disease.
- A person without the disease is correctly classified as not having the disease.
- A person with the disease is misclassified as not having the disease.
- A person without the disease is misclassified as having the disease.

These are the four key features that determine how well a classification test has performed and they can be summarised using three metrics. The accuracy is defined as the percentage of all patients who are correctly classified and is perhaps the most intuitive and general of these metrics. It is can be split into the percentage of patients with the disease who are correctly classified as having the disease (sensitivity) and the percentage of patients without the disease who are correctly classified as not having the disease (specificity).

It is important to look at all three of these metrics, as often a high sensitivity will coincide with a low specificity and vice versa. However, the nature of your question will determine which of these metrics is most important. When interpreting accuracy,

sensitivity and specificity remember that these metrics are not affected by the prevalence of a disease.

Box 6: Diagnostic test example

Consider a diagnostic test where we have results for 100 patients.

		True diagnosis	
		Disease	**No disease**
Diagnostic test	**Disease**	30	5
	No disease	20	45

Accuracy = (30 + 45) /100 = 75%

Sensitivity = 30/50 = 60%

Specificity = 45/50 = 90%

Positive predictive value = 30/35 = 85.7%

Negative predictive value = 45/65 = 69.2%

6.2 Negative predictive value (NPV) and positive predictive values (PPV)

Two further metrics are often used to describe the performance of a classifier: negative predictive value (NPV) and positive predictive value (PPV).

- PPV: the percentage of positive predictions that are true positives
- NPV: the percentage of negative predictions that are true negatives

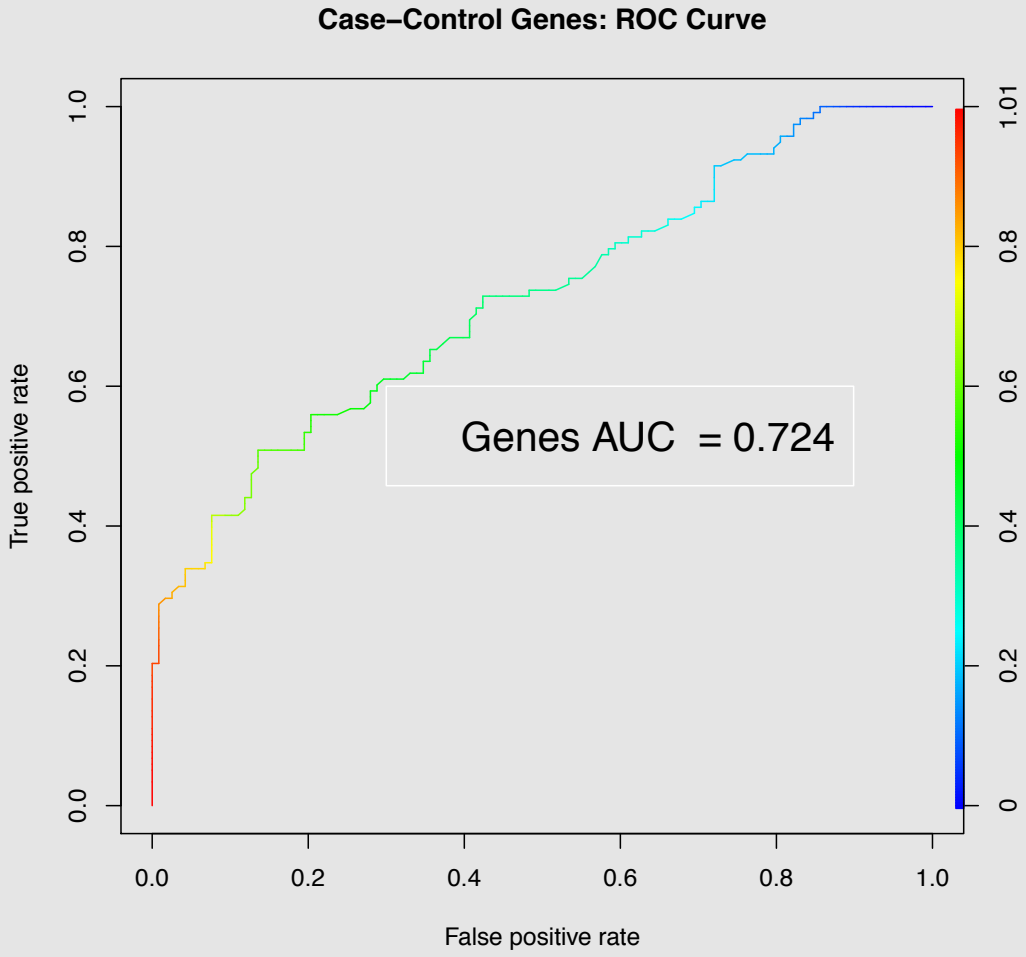
Unlike accuracy, sensitivity and specificity, NPV and PPV will change if the prevalence of a disease changes. PPV will fall as the prevalence of a disease decreases while NPV will rise. This should make sense as a lower prevalence means, over the whole population, a positive result is less likely (Loong, 2003).

6.3 Receiver operating characteristic (ROC) analysis

In most cases, when a classifier is built the underlying method works by creating a probability of assigning a sample to a certain class. If this probability is above a pre-defined threshold then the sample is allocated to the class in question. It is therefore interesting to vary the threshold to see if we can create a test that best answers the question we are studying. A good way to visualise the effect of changing this threshold is through a receiver operating characteristic (ROC) plot. These plots usually have the false positive rate (1 – specificity) on the x-axis and the true positive rate (sensitivity) on the y-axis. They are often summarised by measuring the area under the curve (AUC). A perfect test would have sensitivity and specificity of 1 meaning the AUC would also be 1, as we create a unit square. We therefore look to maximise the AUC to get as close to this perfect test as possible. It is useful to remember that a test that is randomly guessing between two classes should, by chance, achieve an AUC of around 0.5 so any classifier should aim to out-perform this.

Box 7: Example of an ROC curve

Figure 4: An example of an ROC curve generated from a random forest model. The model uses gene expression data to classify subjects as having Alzheimer's disease or being control subjects.



References

Alter, O., Brown, P.O., and Botstein, D. (2000) Singular value decomposition for genome-wide expression data processing and modeling. *PNAS* **97**(18), 10101-10106.

Altman, N. S. (1992) An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician* **46** (3), 175–185.

Barandela, R., Sánchez, J.S., García, V. and Rangel, E. (2003) Strategies for learning in class imbalance problems. *Pattern Recognition* **36**(3), 849-851.

Breiman, L. (2001) Random forests. *Machine Learning* **45**(1), 5-32.

Cortes, C. and Vapnik, V. (1995) Support-Vector Networks. *Machine Learning* **20**(3), 273-297.

Efron, B. and Tibshirani, R. (1994) An introduction to the bootstrap. CRC press.

Fan, W., and Bifet, A. (2013) Mining Big Data : Current Status , and Forecast to the Future. *ACM SIGKDD Explorations Newsletter* **14**(2), 1–5.

Fawcett, T. and Provost, F. (1997) Adaptive Fraud Detection. *Data Mining and Knowledge Discovery* **1**(3), 291-316.

Guyon, I. and Elisseeff, A. (2003) An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research* **3**, 1157-1182.

Hand, D. J., Mannila, H., and Smyth, P. (2001) Principles of data mining. MIT press.

Jain, A.K. (2010) Data clustering: 50 years beyond K-means. *Pattern Recognition Letters* **31**(8), 651-666.

Johnson, S. (2006) The Ghost Map. Riverhead Books, US.

Jolliffe, I. (2014) Principal Component Analysis. Wiley StatsRef: Statistics Reference Online.

Kohavi, R. (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the 14th international joint conference on Artificial intelligence* **2**, 1137-1143.

Laney, D. (2001) 3D Data Management: Controlling Data Volume, Velocity, and Variety. *Gartner **949***(4).

Leek, J. (2015) The Elements of Data Analytic Style.

Loong, T. W. (2003) Understanding sensitivity and specificity with the right side of the brain. *British Medical Journal* **327**, 716-719.

Marsland, S. (2014) Machine Learning. Boca Raton: Chapman & Hall/CRC.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.Y., and Podsakoff, N. P. (2003) Common method biases in behavioral research: a critical review of the literature and recommended remedies. *The Journal of Applied Psychology* **88**(5), 879–903.

Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A. and Reich, D (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics* **38**, 904-909.

Quinlan, J.R. (1993) C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers.

Rousseeuw, P.J. (1987) Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics* **20**, 53-65.

Sayad, S. (2011) Real Time Data Mining. Cambridge, Ontario: Self-Help Publishers.

Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological Methods* **7**(2), 147-177.

Spearman, C. (1904) 'General Intelligence,' Objectively Determined and Measured. *The American Journal of Psychology* **15**(2), 201-292.

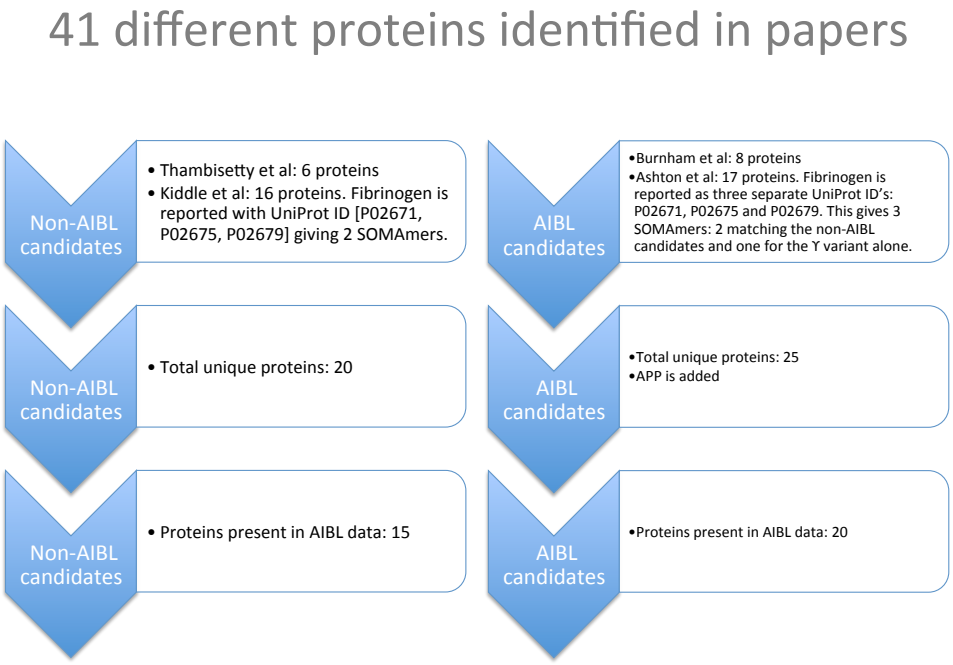
Ward, J.H. (1963) Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association* **58**(301), 236-244.

Wold, H. (2004) Partial least squares. John Wiley & Sons, Inc.

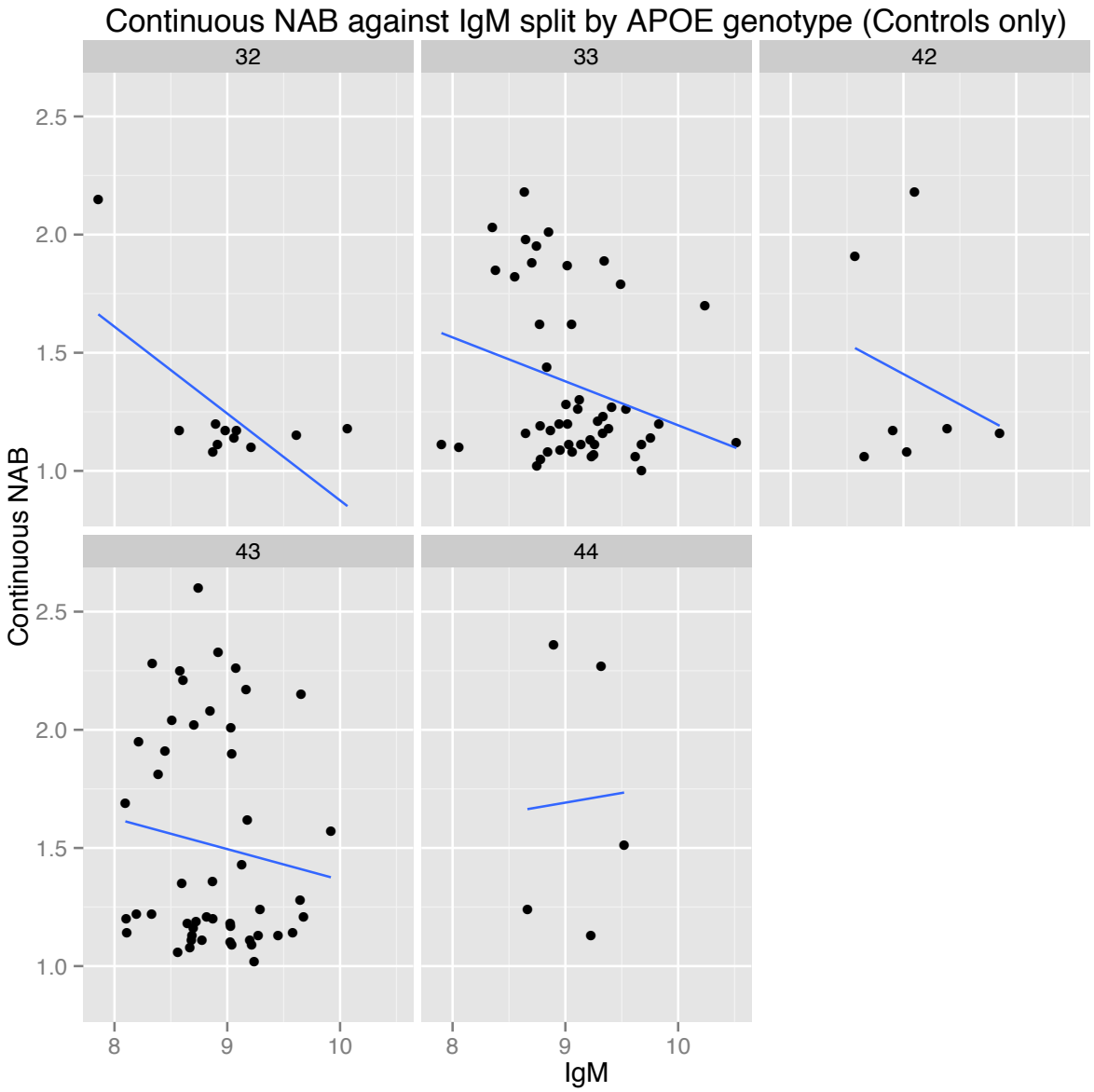
Appendix C

Chapter 2 - Supplementary Material

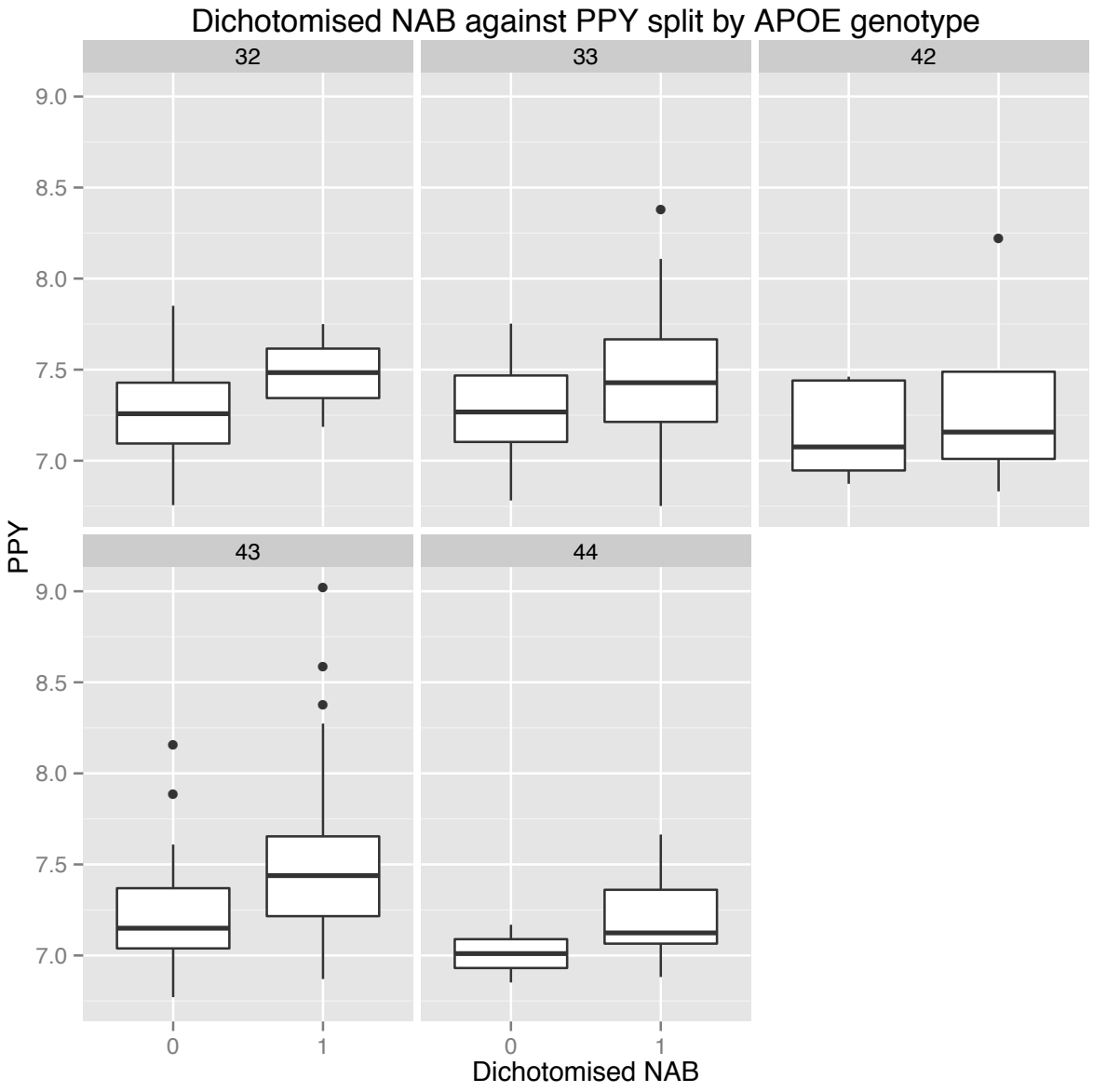
Supplementary Figures



Supplementary Figure 1: 41 different proteins identified in papers



Supplementary Figure 2: Continuous NAB against IgM split by *APOE* genotype (Controls only)



Supplementary Figure 3: Dichotomised NAB against PPY split by *APOE* genotype

Supplementary Tables

Supplementary Table 1: Control population demographics

	Total	**P-value**	**High Aβ burden**	**Low Aβ burden**	**P-value**
	N = 120		N = 43	N = 77	
Gender (% female)	49.2	0.488	41.9	53.2	0.258
APOE status (% of *APOE* ϵ4 positive)	50.8	0.029	62.8	44.2	0.059
APOE load		0.017			0.094
% with load 0	49.2		37.2	55.8	
% with load 1	46.7		55.8	41.6	
% with load 2	4.2		7.0	2.6	
Median age [IQR] (years)	70 [13]	<0.001	75 [12.5]	68 [11]	<0.001
Median MMSE score [IQR]	29 [2]	0.966	29 [2]	29 [2]	0.744
Global CDR status (% ¿ 0)	7.5	0.341	8.3	6.5	0.720

Individuals are positive for *APOE* ϵ4 if at least one *APOE* ϵ4 allele is seen in their genotype.

APOE ϵ4 load is the number of ϵ4 alleles seen in a subjects genotype.

IQR = Inter Quartile Range.

Overall p-value is a result of the Kendall tau test for dependence between SUVR and the demographics variable.

Kruskal Wallis Chi-Squared was used to test between high and low groups for continuous data.

Fishers exact was used to test between high and low groups for categorical data.

Supplementary Table 2: Discovery analysis (included on disc).

Supplementary Table 3: Non-AIBL candidates single protein analysis (included on disc).

Supplementary Table 4: AIBL candidates single protein analysis (included on disc).

Appendix D

Chapter 3 - Supplementary Material

Supplementary Methods

Section 1: PLAGE in detail (Tomfohr *et al.*, 2005)

1. The expression profile of each gene is bought to a common scale by ensuring a distribution of mean zero and variance one.

2. For each pathway a matrix Y is created only containing gene expression data from the genes included in the pathway. In the matrix Y each column represents a sample and each row represents a gene.

3. We perform the singular value decomposition of matrix Y. This involves writing Y in the form:

   Y =WDC

   W = Matrix of eigenvectors ordered by size of corresponding eigenvalue. Each column is an eigenvector.

   D = Diagonal matrix of eigenvalues ordered from largest to smallest.

   C = Matrix of weights. Each column is a vector of coefficients for one sample indicating the overall level of each pathway.

4. The pathway scores for one pathway in all samples are given by the first row of C. The row corresponding to the largest eigenvalue and hence explaining the most variation in the pathway across the samples.

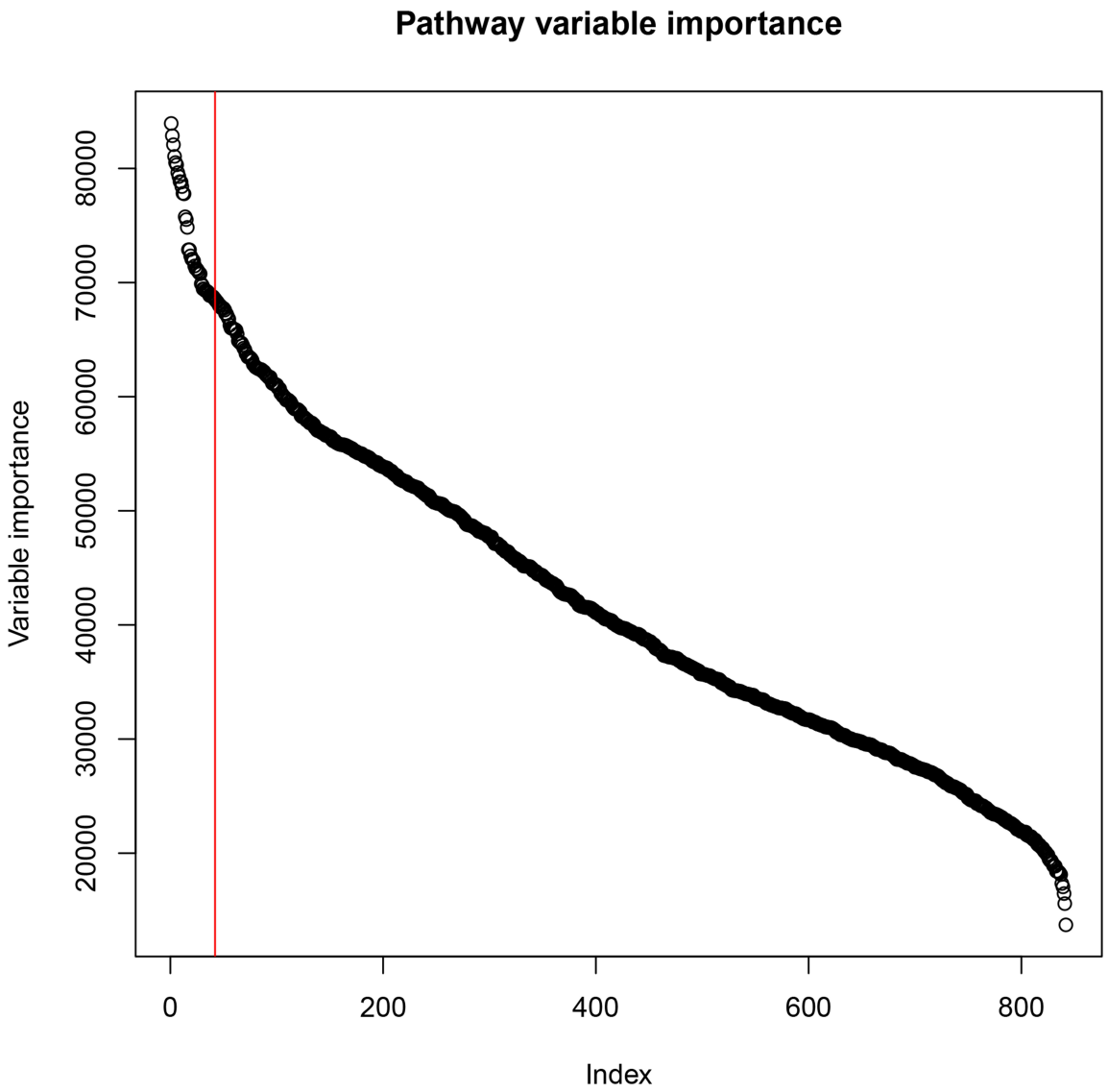
Section 2: Data analysis in detail

To build the gene model and pathway model the batch 1 data was bootstrapped 100 times. The bootstrap samples were of the same size as the batch 1 data and sampled with replacement. For each bootstrap sample a RF model was built using 5 fold cross validation to tune the model parameter mtry (Liaw & Wiener, 2002). mtry is the number of variables randomly selected for consideration at each split in the decision tree. The model parameter that determined the number of trees created within each model ntree was set to 501 throughout; an odd ntree was used to account for any ties. The change in Gini index was used to create variable importance scores. These were ranked across all variables per model and then summed across all bootstrap samples. Variables were ordered by this metric and plotted. The variable importance plot plateaued after approximately 5% of variables and consequently the top 5% of variables were taken forward to the next stage (See Supplementary Figures 1 and 2).

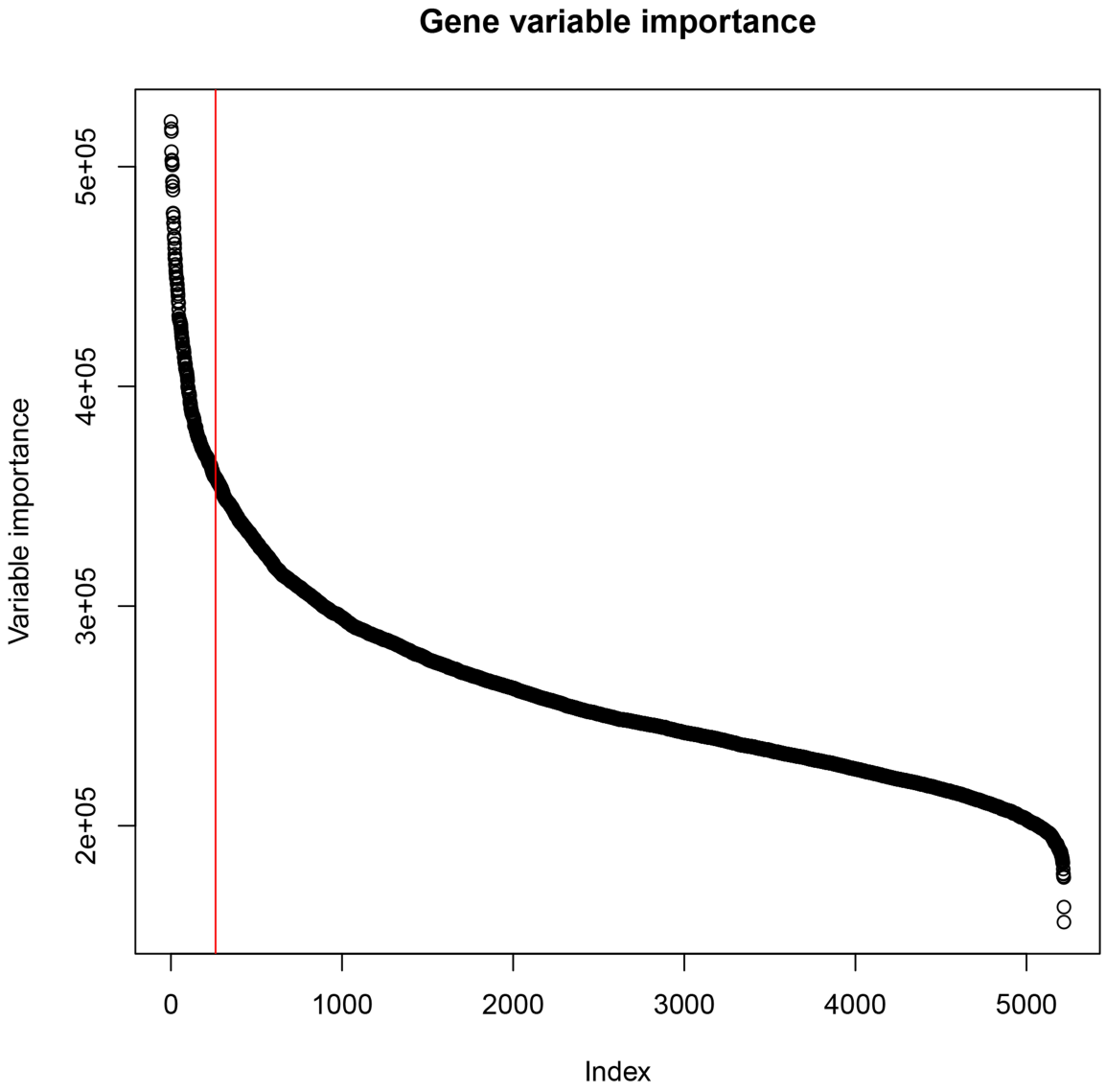
Recursive feature elimination (RFE) was performed on this subset of variables in the original batch 1 data for the pathway model and gene model. Building of the demographic model began at this stage using all variables. The feature elimination was again based on a RF model with ntree = 501 and investigated subsets of variables of all possible sizes. Carets pickSizeTolerance function was used (tolerance = 5%) to identify a further subset of variables. This function finds a smaller set of variables while maintaining model accuracy (Kuhn *et al.*, 2015). If this subset matched the RFE optimal set, the RFE model was taken forward. Otherwise, the optimum variables were selected using the selectVar function in caret and a final RF model was built.

RF models were used throughout for their non-parametric, non-linear properties. Further, the use of bootstrapping in RF modelling and random selection of variables at each decision point decreases the dependence of these models on noise.

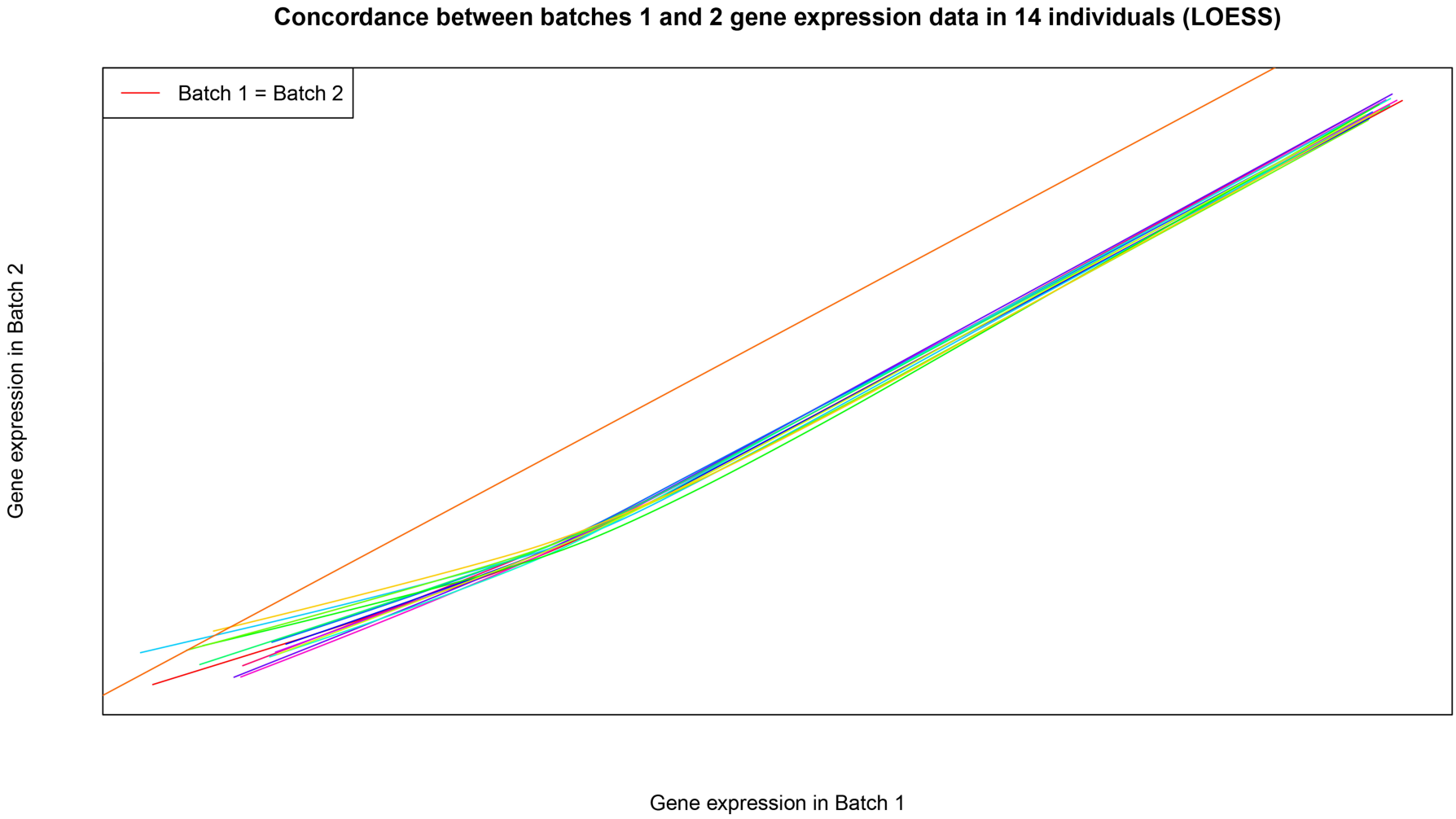
Supplementary Figures



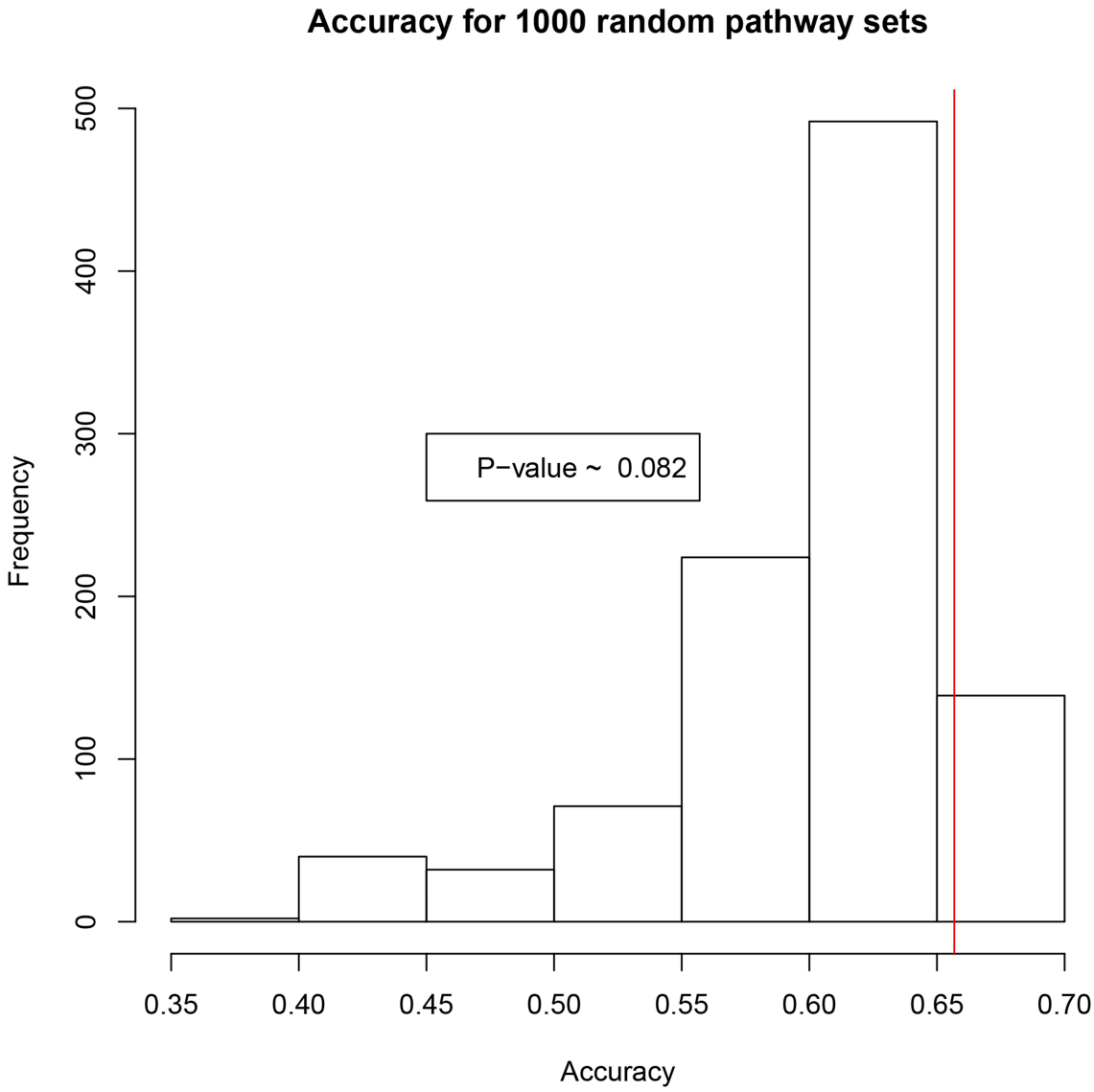
Supplementary Figure 1: Pathway variable importance



Supplementary Figure 2: Gene variable importance

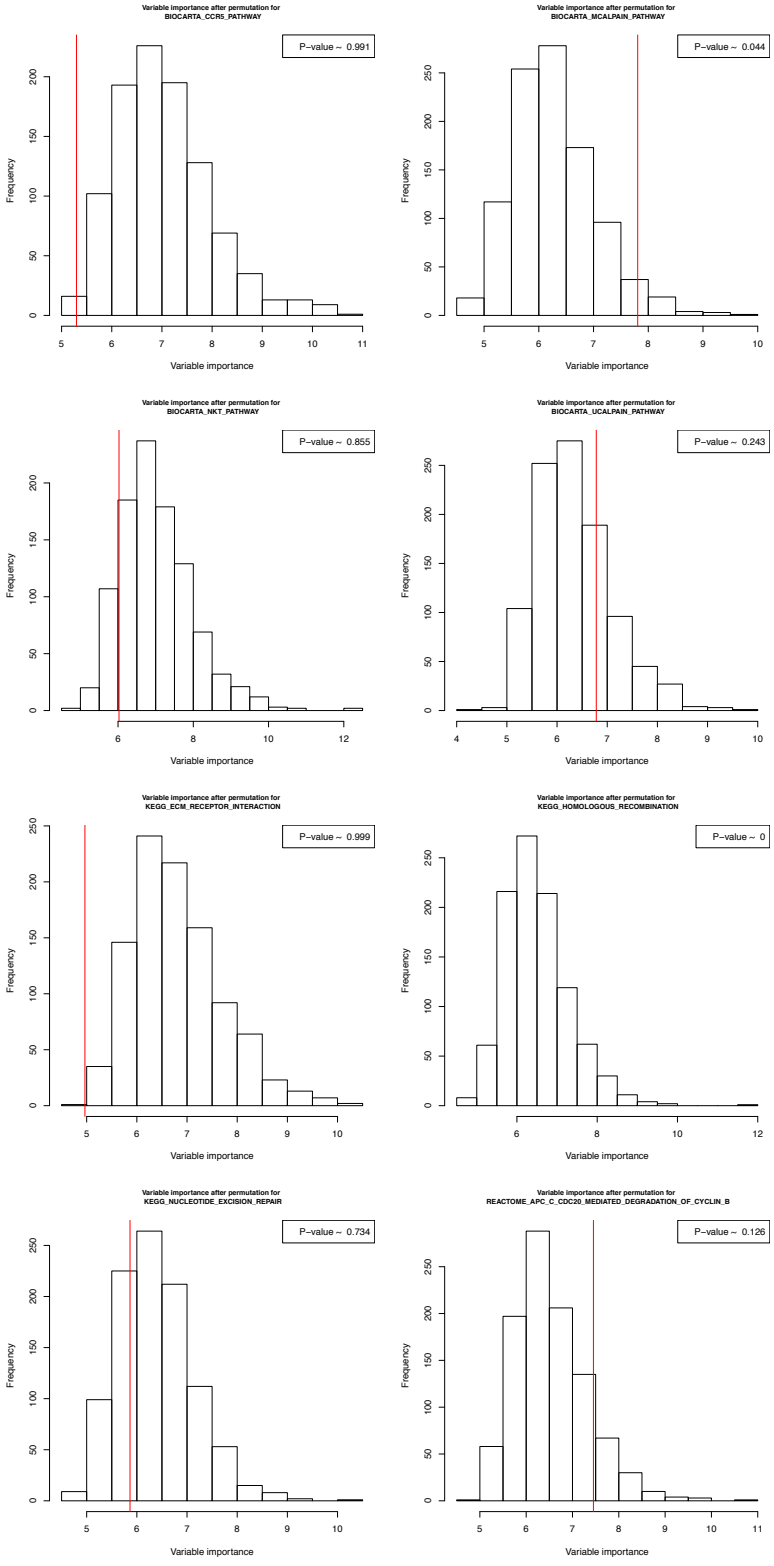


Supplementary Figure 3: Concordance between batches 1 and 2 gene expression data in 14 individuals (LOESS)

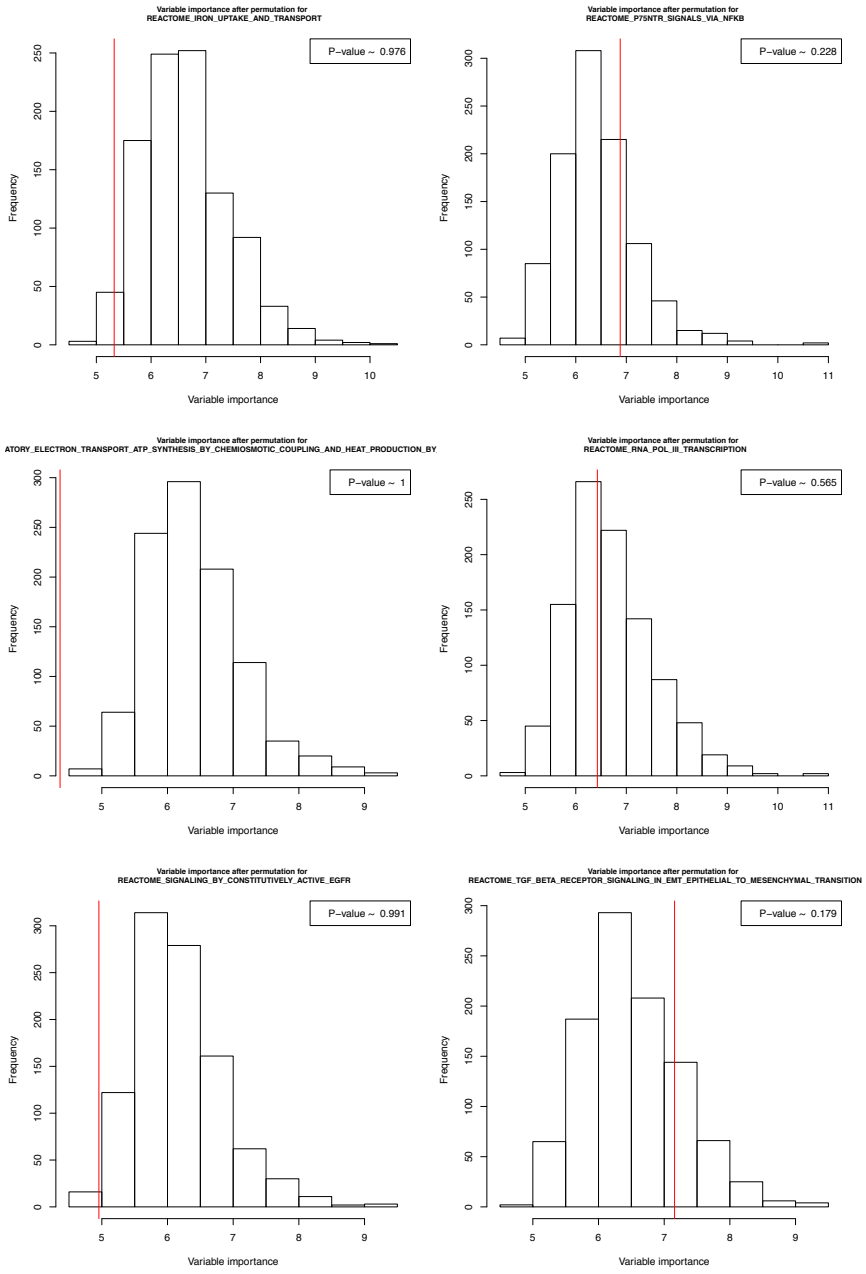


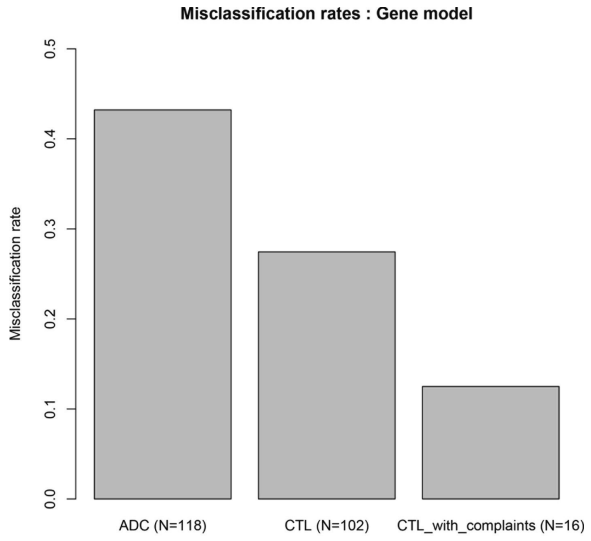
Supplementary Figure 4: Accuracy for 1000 random pathway sets

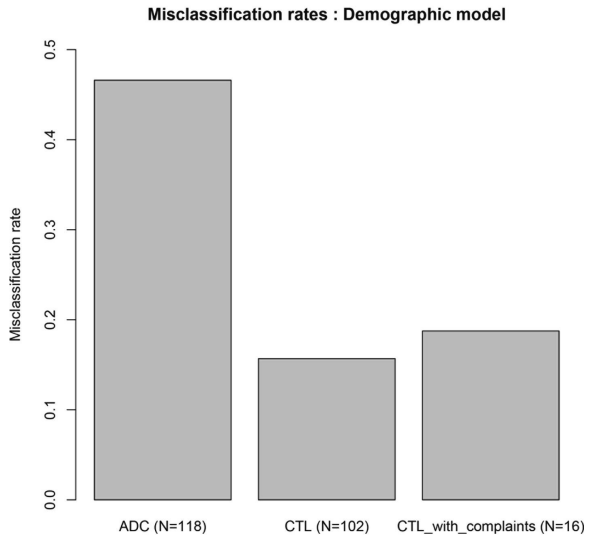
Supplementary Figure 5: Permutation testing of variable importance measures for selected pathways

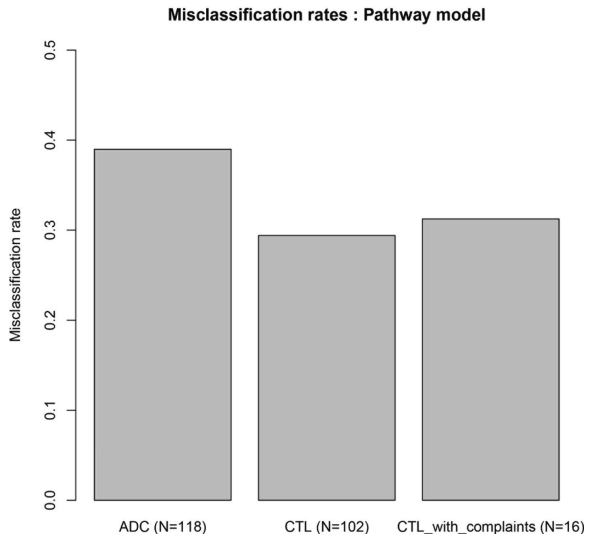


Supplementary Figure 5 continued: Permutation testing of variable importance measures for selected pathways

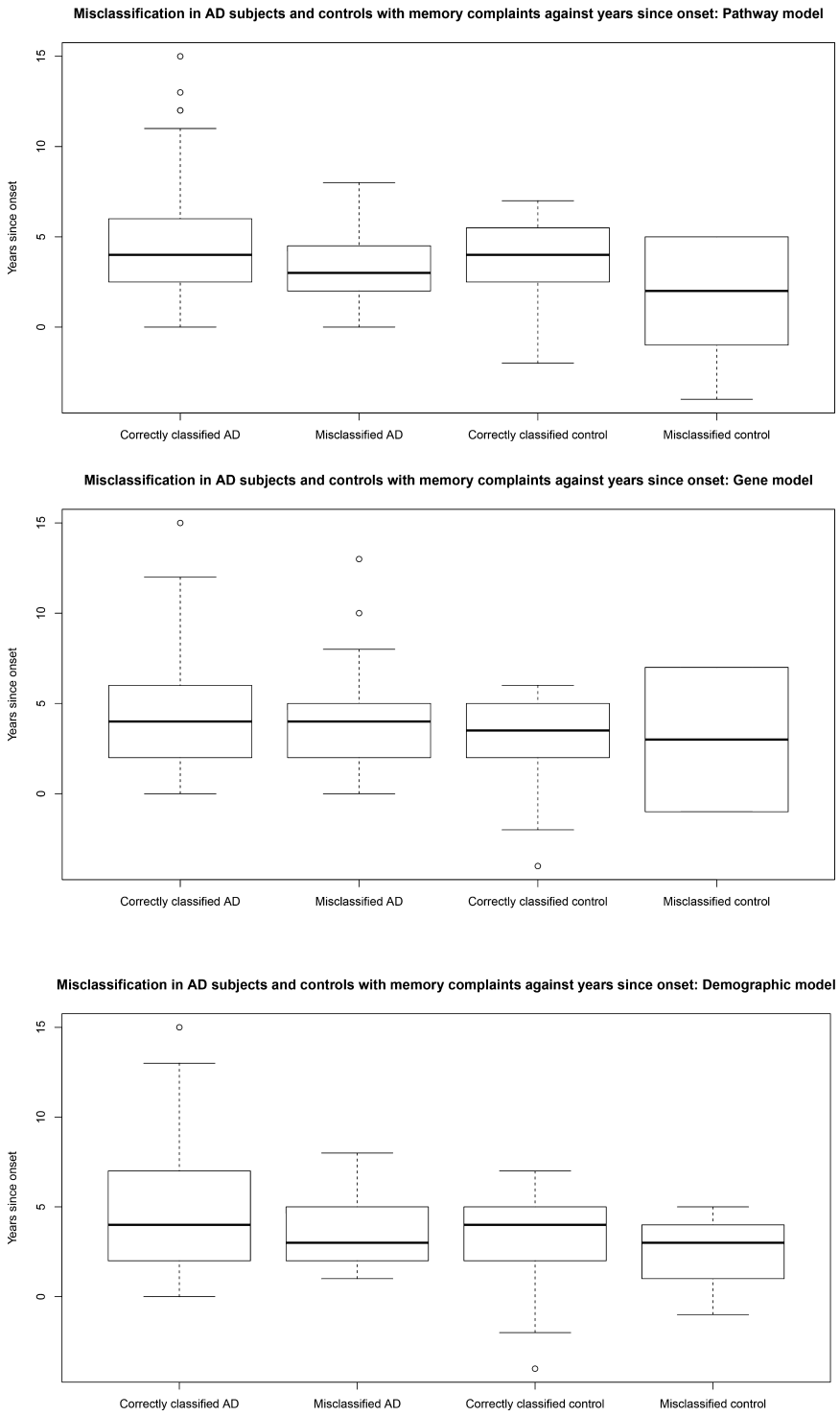








Supplementary Figure 6: Misclassification rates



Supplementary Figure 7: Misclassification in AD subjects and controls with memory complaints against years since onset

Appendix D Bibliography

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., R Core Team, Benesty, M., Lescarbeau, R., Ziem, A. & Scrucca, L. (2015). caret: Classification and Regression Training.

Liaw, A. & Wiener, M. (2002). Classification and Regression by randomForest. *R news*, **2**, 18–22.

Tomfohr, J., Lu, J. & Kepler, T.B. (2005). Pathway level analysis of gene expression using singular value decomposition. *BMC Bioinformatics*, **6**, 211–225.

Appendix E

Chapter 4 - Supplementary Material

Supplementary Methods

Pre-processing of metabolite data

The following pipeline was used to process the metabolic feature data once it had been extracted from netCDF files using the R package 'XCMS' (Smith *et al.*, 2006).

1. Metabolic features that eluted before 1 minute or after 35 minutes were removed.
2. Per sample, per metabolic feature outliers were identified as values outside of 6 standard deviations of the mean metabolic feature value and set to missing.
3. Metabolic features not present in at least 80% of samples were removed.
4. Samples that did not have data from at least 80% of metabolic features were removed.
5. Missingness was investigated per batch. In this study batches were the groups of samples split by QC samples.
6. Each metabolic feature was autoscaled (by subtracting the mean value and dividing by the standard deviation) as suggested by van den Berg *et al.* (2006).
7. In order to remove any negative values created by autoscaling (necessary for the remaining processing steps) each value was increased by 6.
8. The distribution of each metabolic feature was tested for normality using the Shapiro-Wilks test.
9. The data was subject to a log base 2 transformation.
10. Batch effects were removed using ComBat (Johnson *et al.*, 2007).
11. Probabilistic principal components analysis (pPCA) was performed to ensure QC samples clustered and to identify any outlier samples.
12. Missing data was imputed using 10 nearest neighbors.

Supplementary Tables

Supplementary Table 1: Continuous NAB cohort demographics

	Total	**P-value**
	N = 76	
Median NAB SUVR [IQR]	1.3 [0.9]	-
Plasma sample median days in storage [IQR]	1354.5 [568]	0.454
Median number of days difference between sample collection and scan [IQR]	18.5 [73]	0.376
Median age [IQR]	64.91 [10.79]	0.939
Median MMSE [IQR]	26 [6]	< 0.001
Scanner type (%)		0.736
Biograph	9 (11.8)	
ECAT	67 (88.2)	
Gender (%)		0.112
Female	31 (40.8)	
Male	45 (59.2)	
APOE ϵ4 status (%)		0.019
0	57 (75)	
1	19 (25)	
Diagnosis (%)		< 0.001
AD	23 (30.3)	
FTD	47 (61.8)	
HC	4 (5.3)	
MCI	2 (2.6)	

AD = Alzheimers Disease; FTD = Fronto-temporal dementia; HC = Healthy controls; MCI = Mild cognitive impairment.

IQR = Inter-quartile Range; MMSE = Mini Mental State Exam; NAB = Neocortical Amyloid Burden.

SUVR = Standardized Uptake Value Ratio.

P-value: result of the Kendall tau test for dependance between SUVR and the demographics variable.

Supplementary Table 2: Continuous NAB single metabolite results (included on disc).

Supplementary Table 3: Dichotomous NAB single metabolite results (included on disc).

Appendix E Bibliography

Johnson, W.E., Li, C. & Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*, **8**, 118–127.

Smith, C.A., Want, E.J., O'Maille, G., Abagyan, R. & Siuzdak, G. (2006). XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification. *Analytical Chemistry*, **78**, 779–787.

van den Berg, R.A., Hoefsloot, H.C.J., Westerhuis, J.A., Smilde, A.K. & van der Werf, M.J. (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC genomics*, **7**, 142.

Appendix F

Chapter 5 - Supplementary Material

Supplementary Methods

Candidate metabolites (ADNI 1 only)

Metabolite data was available for 853 blood serum samples in ADNI 1. Twenty-four individuals had two samples included in the study. Targeted metabolomic analysis was performed using the AbsoluteIDQ p180 assay requiring 10μl of serum per sample. The samples were run in 11 batches with two pooled QC samples present in each batch: one run before the samples and one afterwards. The metabolite data was processed according to the pipeline described by Voyle *et al.* (2016b). In order to reduce multiple testing burden and over-fitting, one candidate metabolite is used in this study: PCaa36.6 was identified by Voyle *et al.* (2016b) in the only study, to date, investigating associations between metabolites and Aβ. The study identified a panel of 5 metabolites that predicted dichotomised Aβ but PCaa36.6 is the only one also identifiable in ADNI.

Candidate protein assays (ADNI 1 only)

ADNI 1 blood proteomics data is provided by the Biomarkers Consortium Plasma Proteomics Project for Rules-Based Medicine (RBM) multiplex data. In short, blood plasma samples were analysed on the human discovery map; a 190 analyte immunoassay panel developed by RBM. The panel was designed to include plasma proteins thought to be involved in cancer, cardiovascular disease, metabolic disorders, inflammation and AD (Ray *et al.*, 2007). Samples were selected for proteomic analysis based on additional biomarker endpoints being available. In particular, control samples were chosen to have baseline CSF Aβ levels above the median of all control participants. This study considers one candidate protein that has shown some replicability in other proteomic studies: pancreatic polypeptide (PPY) as reviewed in Voyle *et al.* (2015) (Burnham *et al.*, 2013; Kiddle *et al.*, 2012). Fibrinogen gamma chain (FGG) has also shown some replication (Ashton *et al.*, 2015) but the RBM probe for FGG is known to also bind to alpha and beta chains so is not used in this analysis (Kiddle *et al.*, 2012).

The data was checked for outliers, defined as values greater than 6 standard deviations from

the mean, and none were found. The data was approximately normally distributed (Shaprio-Wilks test p-value = 0.67) so no data transformation was performed.

Gene expression

EDAR

The EDAR gene expression data (N = 115) was generated on a Illumina HT-12 v4.0 expression beadchip and processed using the pipeline previously used by Voyle *et al.* (2016a) (http://bit.ly/1vjyKNo). In short, raw expression data was subject to a model based background correction for bead array (Ding *et al.*, 2008). The data was then log base 2 transformed and robust spline normalized. Outlying samples were iteratively identified using fundamental network concepts and removed (Oldham *et al.*, 2012). Technical artefacts were accounted for by principal components analysis and the use of ComBat (Johnson *et al.*, 2007). Finally, the data was subset to probes that could be reliably detected in at least 80% of samples in at least one diagnostic group. After processing gene expression data was available for 109 individuals.

ADNI 2

Gene expression data was collected for 811 ADNI subjects on the Affymetrix Human Genome U219 Array. The array contains 530,467 probes for 49,293 transcripts mapped and annotated with reference to the human genome (hg19). The publicly available data has been processed using robust multi-chip averaging (RMA) and gender checked (Irizarry *et al.*, 2003). Additional processing was performed to ensure the data was as similar as possible to that from EDAR, despite the differences between Affymetrix and Illumina arrays. The processing steps were:

1. Identification of non-expressed probes: probes whose mean expression level was in the lowest 20th percentile for all diagnostic groups.
2. Removal of outlier samples using network concepts (Oldham *et al.*, 2012).
3. Adjustment for continuous batch effects using linear regression models (RIN and 260\230

ratio of nucleic acid purity) and discrete batch effect using ComBat (plate and sex) (Johnson *et al.*, 2007).

4. Removal of probes identified as non-expressed in step 1.

This pipeline is adapted from http://bit.ly/1vjyKNo to be compatible with an Affymetrix array.

Gene Expression Risk Score (GERS)

GERS were calculated using a method based on that presented by Bret *et al.* (2012) for all 109 individuals from the EDAR study with gene expression data. In short, all expression probes were regressed against Aβ burden using a LASSO regression model. This method selects a subset of probes by shrinking the estimates of non-informative variables to zero. The regularization parameter (lambda) in the LASSO regression was tuned using 10 fold cross-validation (CV). In order to choose the simplest informative model we selected the largest lambda that gave a CV error within 1 standard deviation of the minimum CV error. The subset of probes were then tested for prognostic significance using the Wilcoxon rank statistic through the R function 'Maxstat' (Hothorn, 2015). This produced a gene expression cut-off that best differentiated between high and low Aβ burden for each probe. The cut-off was converted to a rank based metric to enable application to other datasets, possibly generated on alternative chips. The per subject GERS was calculated by summing the coefficients from the LASSO regression weighted by +1 if a signal was greater than the cut-off value defined by maxstat, and -1 if the signal was less than or equal to the cut-off. The scores were built in the EDAR cohort and the estimates, and rank based cut-offs, applied to create the GERS in ADNI 2. Supplementary Figure 1 outlines the steps used to create both the GERS and the PGRS.

In this study, the LASSO regression used to generate the GERS maintained non-zero estimates for 25 probes. More details of these probes can be found in Supplementary Table 5.

Modelling

Analysis 1

Models of dichotomized Aβ burden were built using 5 fold cross-validation in ADNI 1 (N = 222). A standardized PGRS and measurements of PPY and PC.aa.36.6 were included in the modelling. Demographics of the population are given in Supplementary Table 1 and results in Supplementary Table 3.

Analysis 2

Models of dichotomized Aβ burden were built in EDAR (N = 47) and tested in ADNI 2 (N = 38). A standardized PGRS and a GERS were included in the modelling. Demographics of the population are given in Supplementary Table 2 and results in Supplementary Table 4.

Supplementary Tables

Supplementary Table 1: ADNI 1 cohort demographics ANALYSIS 1

	Overall (N=222)	**Normal CSF Aβ (N=95)**	**Abnormal CSF Aβ (N=127)**	**P-value**
Median age [IQR]	74.15 [8.7]	74 [8.45]	74.7 [9.75]	0.816
Gender (%)				
Female	35.6	33.7	37	0.672
Male	64.4	66.3	63	
Median years in education [IQR]	16 [4]	16 [4.5]	16 [4]	0.821
Median MMSE [IQR]	28 [3]	29 [2.5]	27 [3]	< 0.001
Diagnosis (%)				
CTL	22.5	52.6	0	< 0.001
MCI	77.5	47.4	100	
APOE (%)				
0	54.5	84.2	32.3	< 0.001
1	45.5	15.8	67.7	
Standardized PGRS [IQR]	-0.567 [2.150]	-0.595 [2.214]	-0.557 [0.157]	0.01
PC.aa.36.6 [IQR]	2.534 [0.281]	2.541 [0.253]	2.530 [0.337]	0.868
PPY [IQR]	2.055 [0.480]	2.045 [0.441]	2.072 [0.503]	0.114

Kruskal Wallis chi-squared was used to test between normal and abnormal groups for continuous demographic variables.

Fishers exact was used to test between normal and abnormal groups for categorical demographic variables.

APOE status is 1 if an individuals genotype contains any e4 alleles, and 0 otherwise.

IQR = Inter-quartile range; CSF = Cerebrospinal Fluid; MMSE = Mini Mental State Exam;

MCI = Mild Cognitive Impairment; CTL = Control; PGRS = Polygenic risk score; PPY = Pancreatic Polypeptide.

Supplementary Table 2: Cohort demographics ANALYSIS 2

EDAR				
	Overall (N=47)	**Normal CSF Aβ (N=24)**	**Abnormal CSF Aβ (N=23)**	**P-value**
Median age [IQR]	69 [10.5]	65 [11]	69 [10.5]	0.088
Gender (%)				
Female	36.2	29.2	43.5	0.371
Male	63.8	70.8	56.5	
Median years in education [IQR]	12 [6.5]	14 [4.25]	11 [6.5]	0.08
Median MMSE [IQR]	27 [3]	28 [4]	27 [3]	0.203
Diagnosis (%)				
CTL	10.6	12.5	8.7	> 0.999
MCI	89.4	87.5	91.3	
APOE (%)				
0	44.7	58.3	30.4	0.08
1	55.3	41.7	69.6	
Standardized PGRS [IQR]	0.142 [1.689]	-0.058 [1.341]	0.484 [1.380]	0.202
GERS [IQR]	0.184 [5.679]	-2.883 [2.660]	3.025 [3.211]	< 0.001

ADNI 2				
	Overall (N=38)	**Normal CSF Aβ (N=22)**	**Abnormal CSF Aβ (N=16)**	**P-value**
Median age [IQR]	69.5 [11.3]	67.7 [11.5]	73.75 [11.1]	0.048
Gender (%)				
Female	50	59.1	37.5	0.325
Male	50	40.9	62.5	
Median years in education [IQR]	16 [4]	16 [2.75]	16 [4]	0.553
Median MMSE [IQR]	29 [2.75]	29 [2]	27.5 [2.25]	0.103
Diagnosis (%)				
MCI	81.6	63.6	87.5	0.675
CTL	18.4	36.4	12.5	
APOE (%)				
0	57.9	63.6	50	0.511
1	42.1	36.4	50	
Standardized PGRS [IQR]	0.0002 [1.246]	0.089 [1.323]	-0.221 [1.127]	0.554
GERS [IQR]	-0.821 [3.968]	-1.297 [3.999]	0.078 [3.429]	0.117

Kruskal Wallis chi-squared was used to test between normal and abnormal groups for continuous demographic variables.

Fishers exact was used to test between normal and abnormal groups for categorical demographic variables.

APOE status is 1 if an individuals genotype contains any e4 alleles, and 0 otherwise.

IQR = Inter-quartile range; CSF = Cerebrospinal Fluid; MMSE = Mini Mental State Exam;

MCI = Mild Cognitive Impairment; CTL = Control; PGRS = Polygenic risk score; GERS = Gene expression risk score.

Supplementary Table 3: Results for ANALYSIS 1

Model	Informative priors?	Accuracy	Sensitivity	Specificity	AUC ROC
Demographics	NO	0.6576	0.7148	0.5926	0.767
Demographics	YES	0.5182	0.5362	0.4896	0.5312
Demo + Protein	NO	0.387	0.5434	0.1852	0.4368
Demo + Protein	YES	0.4956	0.1414	0.9676	0.4344
Demo + Metabolite	NO	0.3692	0.5434	0.142	0.3398
Demo + Metabolite	YES	0.4864	0.1244	0.9676	0.4328
Demo + PGRS	NO	0.4598	0.252	0.7316	0.4198
Demo + PGRS	YES	0.441	0.5162	0.3614	0.4626
Demo + Protein + Metabolite	NO	0.4282	0	1	0.5686
Demo + Protein + Metabolite	YES	0.4282	0	1	0.656
Demo + Protein + PGRS	NO	0.4282	0	1	0.4224
Demo + Protein + PGRS	YES	0.518	0.2592	0.8656	0.6398
Demo + Metabolite + PGRS	NO	0.5714	0.7708	0.3162	0.6346
Demo + Metabolite + PGRS	YES	0.4282	0	1	0.6078
Demo + Protein + Metabolite + PGRS	NO	0.5714	0.7708	0.3162	0.4912
Demo + Protein + Metabolite + PGRS	YES	0.4282	0	1	0.5264

PGRS = Polygenic Risk Score

AUC ROC = Area under the Receiver Operating Characteristic Curve

Supplementary Table 4: Results for ANALYSIS 2

Model	Informative priors?	Accuracy	Sensitivity	Specificity	AUC ROC
Demographics	NO	0.421 [0.263; 0.592]	0.625	0.273	0.315
Demographics	YES	0.5 [0.334; 0.666]	0.625	0.409	0.415
Demo + PGRS	NO	0.579 [0.408; 0.737]	0.625	0.545	0.661
Demo + PGRS	YES	0.395 [0.24; 0.566]	0.438	0.364	0.347
Demo + GERS	NO	0.474 [0.31; 0.642]	0.688	0.318	0.472
Demo + GERS	YES	0.342 [0.196; 0.514]	0.375	0.318	0.27
Demo + PGRS + GERS	NO	0.579 [0.408; 0.737]	0.688	0.5	0.636
Demo + PGRS + GERS	YES	0.474 [0.31; 0.642]	0.562	0.409	0.42

AUC ROC = Area under the Receiver Operating Characteristic Curve.

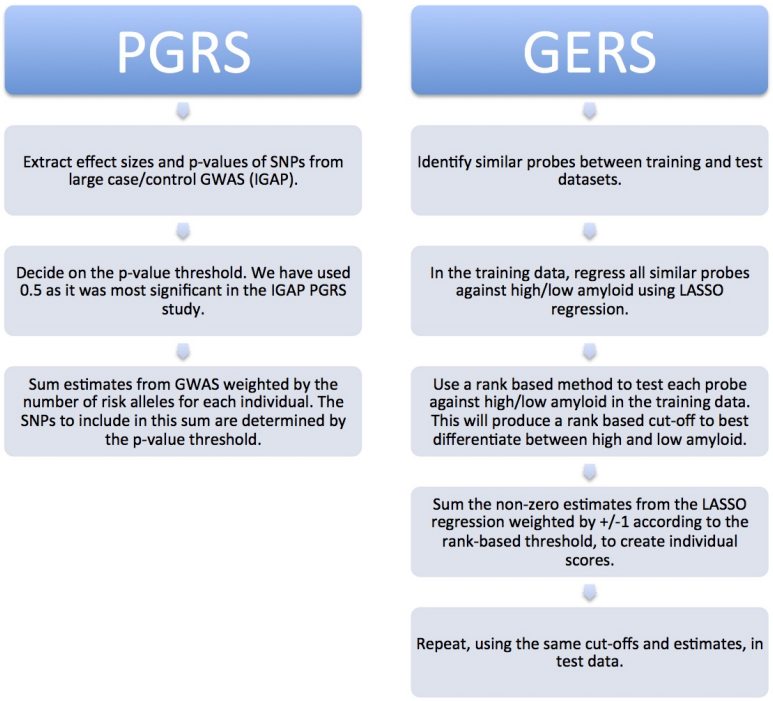
PGRS = Polygenic Risk Score; GERS = Gene expression risk score.

Supplementary Table 5: Probes in GERS in ANALYSIS 2

Entrez ID	LASSO coefficient (3dp)	Gene name
10077	0.053	*TSPAN32*: Tetraspanin 32
10425	0.01	*ARIH2*: AriadneRBR E3 ubiquitin protein ligase 2
10961	2.297	*ERP29*: Endoplasmic reticulum protein 29
1861	0.003	*TOR1A*: Torsin family 1, member A
212	-0.273	*ALAS2*: 5'-aminolevulinate synthase 2
3126	0.022	*HLA-DRB4*: Major histocompatibility complex, class II, DR beta 4
3543	-0.177	*IGLL1*: Immunoglobulin lambda-like polypeptide 1
4069	0.321	*LYZ*: Lysozyme
51024	-0.203	*FIS1*: Fission, mitochondrial 1
51312	-0.09	*SLC25A37*: Solute carrier family 25 (mitochondrial iron transporter), member 37
51399	0.873	*TRAPPC4*: Trafficking protein particle complex 4
54807	0.043	*ZNF586*: Zinc finger protein 586
55766	0.489	*H2AFJ*: H2A histone family member J
57222	-0.136	*ERGIC1*: Endoplasmic reticulum-golgi intermediate compartment 1
6228	0.255	*RPS23*: Ribosomal protein S23
65264	-0.078	*UBE2Z*: ubiquitin conjugating enzyme E2Z
80162	0.448	*ATHL1*: Acide trehalase-like 1
8349	0.553	*HIST2H2BE*: Histone cluster 2
84273	0.035	*NOA1*: Nitric oxide associated 1
84447	-0.105	*SYVN1*: Synoviolin 1
84545	-0.405	*MRPL43*: Mitochondrial ribosomal protein L43
8546	0.53	*AP3B1*: Adaptor-related protein complex 3, beta 1 subunit
89790	-0.299	*SIGLEC10*: Sialic acid binding Ig-like lectin 10
973	-0.297	*CD79A*
997	-0.036	*CDC34*: Cell division cycle 34

GERS = Gene expression risk score; dp = decimal places.

Supplementary Figures



PGRS = Polygenic Risk Score; GERS = Gene Expression Risk Score; SNP = Single Nucleotide Polymorphism; GWAS = Genome Wide Association Study; IGAP = International Genomics of Alzheimer's Project; LASSO = Least Absolute Shrinkage and Selection Operator.

Supplementary Figure 1: Overview of PGRS and GERS for ANALYSIS 2

Appendix F Bibliography

Ashton, N.J., Kiddle, S.J., Graf, J., Ward, M., Baird, A.L., Hye, A., Westwood, S., Wong, K.V., Dobson, R.J., Rabinovici, G.D., Miller, B.L., Rosen, H.J., Torres, A., Zhang, Z., Thurfjell, L., Covin, A., Hehir, C.T., Baker, D., Bazenet, C. & Lovestone, S. (2015). Blood protein predictors of brain amyloid for enrichment in clinical trials? *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, **1**, 48–60.

Bret, C., Klein, B. & Moreaux, J. (2012). Gene expression-based risk score in diffuse large B-celllymphoma. *Oncotarget*, **3**, 1700–1710.

Burnham, S.C., Faux, N.G., Wilson, W., Laws, S.M., Ames, D., Bedo, J., Bush, A.I., Doecke, J.D., Ellis, K.A., Head, R., Jones, G., Kiiveri, H., Martins, R.N., Rembach, A., Rowe, C.C., Salvado, O., Macaulay, S.L., Masters, C.L. & Villemagne, V.L. (2013). A blood-based predictor for neocortical Aβ burden in Alzheimer's disease: results from the AIBL study. *Molecular Psychiatry*, **19**, 519–526.

Ding, L.H., Xie, Y., Park, S., Xiao, G. & Story, M.D. (2008). Enhanced identification and biological validation of differential gene expression via Illumina whole-genome expression arrays through the use of the model-based background correction methodology. *Nucleic Acids Research*, **36**, e58–e58.

Hothorn, T. (2015). maxstat: Maximally Selected Rank Statistics.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U. & Speed, T.P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*, **4**, 249–264.

Johnson, W.E., Li, C. & Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*, **8**, 118–127.

Kiddle, S.J., Thambisetty, M., Simmons, A., Riddoch-Contreras, J., Hye, A., Westman, E., Pike, I., Ward, M., Johnston, C., Lupton, M.K., Lunnon, K., Soininen, H., Kloszewska, I., Tsolaki, M., Vellas, B., Mecocci, P., Lovestone, S., Newhouse, S., Dobson, R. & for the Alzheimers Disease Neuroimaging Initiative (2012). Plasma Based Markers of [11C] PiB-PET Brain Amyloid Burden. *PLOS ONE*, **7**, e44260.

Oldham, M.C., Langfelder, P. & Horvath, S. (2012). Network methods for describing sample relationships in genomic datasets: application to Huntington's disease. *BMC Systems Biology*, **6**, 1.

Ray, S., Britschgi, M., Herbert, C., Takeda-Uchimura, Y., Boxer, A., Blennow, K., Friedman, L.F., Galasko, D.R., Jutel, M., Karydas, A., Kaye, J.A., Leszek, J., Miller, B.L., Minthon, L., Quinn, J.F., Rabinovici, G.D., Robinson, W.H., Sabbagh, M.N., So, Y.T., Sparks, D.L., Tabaton, M., Tinklenberg, J., Yesavage, J.A., Tibshirani, R. & Wyss-Coray, T. (2007). Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins. *Nature Medicine*, **13**, 1359–1362.

Voyle, N., Baker, D., Burnham, S.C., Covin, A., Zhang, Z.P., Sangurdekar, D.P., Hehir, C.A.T., Bazenet, C., Lovestone, S., Kiddle, S. & Dobson, R.J.B. (2015). Blood Protein Markers of Neocortical Amyloid-beta Burden: A Candidate Study Using SOMAscan Technology. *Journal of Alzheimers Disease*, **46**, 947–961.

Voyle, N., Keohane, A., Newhouse, S., Lunnon, K., Johnston, C., Soininen, H., Kloszewska, I., Mecocci, P., Tsolaki, M., Vellas, B., Lovestone, S., Hodges, A., Kiddle, S. & Dobson, R.J.B. (2016a). A Pathway Based Classification Method for Analyzing Gene Expression for Alzheimer' s Disease Diagnosis. *Journal of Alzheimer's Disease*, **49**, 659–669.

Voyle, N., Kim, M., Proitsi, P., Ashton, N.J., Baird, A.L., Bazenet, C., Hye, A., Westwood, S., Chung, R., Ward, M., Rabinovici, G.D., Lovestone, S., Breen, G., Legido-Quigley, C., Dobson, R.J.B. & Kiddle, S.J. (2016b). Blood metabolite markers of neocortical amyloid-

β burden: discovery and enrichment using candidate proteins. *Translational Psychiatry*, **6**, e719.

Appendix G

Chapter 6 - Supplementary Material

Supplementary Tables

Supplementary Table 1: Cohort demographics against total CSF

Demographic	ADNI 1 (N = 244)			EDAR and DESCRIPA* (N = 135)			ADNI 2 (N = 37)		
	Normal total CSF (N = 98)	Abnormal total CSF (N = 146)	P-value	Normal total CSF (N = 67)	Abnormal total CSF (N = 68)	P-value	Normal total CSF (N = 25)	Abnormal total CSF (N = 12)	P-value
Median age [IQR]	74.2 [8.37]	74.4 [10.25]	0.729	66 [10.5]	70 [12.25]	0.013	67.8 [12.8]	78.35 [13.3]	0.031
Median years in education [IQR]	16 [4]	16 [5]	0.258	12 [7]	8.5 [7]	0.02	16 [4]	15.5 [4]	0.754
Median MMSE [IQR]	29 [2.75]	26 [4]	< 0.001	28 [3]	25 [4.25]	< 0.001	29 [2]	27 [2.75]	0.059
APOE status (%):									
0	84.7	26	< 0.001	71.6	25	< 0.001	68	41.7	0.164
1	15.3	74		28.4	75		32	58.3	
Gender (%):									
Female	34.7	42.5	0.232	38.8	60.3	0.016	48	33.3	0.491
Male	65.3	57.5		61.2	39.7		52	66.7	
Diagnosis (%):									
Dementia	5.1	39	< 0.001	22.3	63.3	< 0.001	0	8.3	0.059
MCI	42.9	52.7		47.8	33.8		80	91.7	
SCI	0	0		7.5	2.9		0	0	
CTL	52	8.2		22.4	0		20	0	

Kruskall Wallis chi-squared was used to test between normal and abnormal groups for continuous demographics variables.

Fisher's exact was used to test between normal and abnormal groups for categorical demographic variables.

APOE status is 1 if an individuals genotype contains any e4 alleles, and 0 otherwise.

IQR = Inter-quartile range; CSF = Cerebrospinal Fluid; MMSE = Mini Mental State Exam; MCI = Mild Cognitive Impairment; SCI = Subjective Cognitive Impairment; CTL = Control.

* One individual has missing diagnosis, 2 have missing education information and 3 have missing MMSE.

Supplementary Table 2: Cohort demographics for ADNI 1 individuals with plasma tau measurements

	ADNI 1 (N = 323)			ADNI 1 (N = 323)			ADNI 1 (N = 219)		
Demographic	**Normal CSF tTau** (N = 180)	**Abnormal CSF tTau** (N = 143)	**P-value**	**Normal CSF amyloid** (N = 98)	**Abnormal CSF amyloid** (N = 225)	**P-value**	**Normal total CSF** (N = 87)	**Abnormal total CSF** (N = 132)	**P-value**
Median age [IQR]	75.2 [8.05]	74.7 [9.7]	0.844	74 [8.73]	75.6 [8.5]	0.723	74 [8.70]	74.6 [10.18]	0.847
Median years in education [IQR]	16 [4]	16 [5]	0.054	16 [5]	16 [4]	0.991	16 [5]	16 [5]	0.332
Median MMSE [IQR]	28 [3]	26 [4]	< 0.001	29 [3]	26 [4]	< 0.001	29 [2.5]	26 [4]	< 0.001
APOE status (%):									
0	62.8	30.8	< 0.001	84.7	32.9	< 0.001	83.9	25.8	< 0.001
1	37.2	69.2		15.3	67.1		16.1	74.2	
Gender (%):									
Female	35	42.7	0.169	34.7	40	0.386	33.3	42.4	0.203
Male	65	57.3		65.3	60		66.7	57.6	
Diagnosis (%):									
Dementia	14.4	39.2		6.1	33.8		5.7	41.7	
MCI	47.3	49.6	< 0.001	39.8	52	< 0.001	42.5	52.3	< 0.001
SCI	0	0		0	0		0	0	
CTL	38.3	11.2		54.1	14.2		51.7	6.1	

Kruskall Wallis chi-squared was used to test between normal and abnormal groups for continuous demographics variables.

Fisher's exact was used to test between normal and abnormal groups for categorical demographic variables.

APOE status is 1 if an individuals genotype contains any e4 alleles, and 0 otherwise.

IQR = Inter-quartile range; CSF = Cerebrospinal Fluid; MMSE = Mini Mental State Exam; MCI = Mild Cognitive Impairment; SCI = Subjective Cognitive Impairment; CTL = Control.